RTO-EN-3

RTO-EN-3
AC/323(HFM)TP/1

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO LECTURE SERIES 215

Alternative Control Technologies: Human Factors Issues

(Techniques de pilotage alternatives – Le facteur humain)

The material in this publication was assembled to support a Lecture Series under the sponsorship of the Human Factors and Medicine Panel and the Consultant and Exchange Programme of RTO presented on 7-8 October 1998 in Brétigny, France, and on 14-15 October 1998 at Wright-Patterson Air Force Base, Ohio, USA.

Published October 1998

Distribution and Availability on Back Cover

North Atlantic Treaty Organization

Research and Technology Agency

RTA Headquarters: 7, rue Ancelle - 92200 Neuilly-sur-Seine, France

ST/60/4 19 August, 1998

TO: Recipients of RTO Publications

FROM: Scientific Publications Executive

SUBJECT: **RTO Technical Publications**

As you probably know, NATO formed the Research and Technology Organization (RTO) on 1 January 1998, by merging the former AGARD (Advisory Group for Aerospace Research and Development) and DRG (Defence Research Group). There is a brief description of RTO on page ii of this publication.

This new organization will continue to publish high-class technical reports, as did the constituent bodies. There will be five series of publications:

AG **AGARDographs** (Advanced **G**uidance for **A**lliance **R**esearch and **D**evelopment), a successor to the former AGARD AGARDograph series of monographs, and containing material of the same long-lasting value.

MP **Meeting Proceedings**: the papers presented at non-educational meetings at which the attendance is not limited to members of RTO bodies. This will include symposia, specialists' meetings and workshops. Some of these publications will include a Technical Evaluation Report of the meeting and edited transcripts of any discussions following the presentations.

EN **Educational Notes**: the papers presented at lecture series or courses.

TR **Technical Reports**: other technical publications given a full distribution throughout the NATO nations (within any limitations due to their classification).

TM **Technical Memoranda**: other technical publications not given a full distribution, for example because they are of ephemeral value only or because the results of the study that produced them may be released only to the nations that participated in it.

The first series (AG) will continue numbering from the AGARD series of the same name, although the publications will now relate to all aspects of defence research and technology and not only aerospace as formerly. The other series will start numbering at 1, although (as in the past) the numbers may not appear consecutively because they are generally allocated about a year before the publication is expected.

All publications, like this one, will also have an 'AC/323' number printed on the cover. This is mainly for use by the NATO authorities.

Please write to me (do not telephone) if you want any further information.

G.W.Hart

RTO-EN-3
AC/323(HFM)TP/1

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION
BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO LECTURE SERIES 215

Alternative Control Technologies: Human Factors Issues

(Techniques de pilotage alternatives – Le facteur humain)

The material in this publication was assembled to support a Lecture Series under the sponsorship of the Human Factors and Medecine Panel and the Consultant and Exchange Programme of RTO presented on 7-8 October 1998 in Brétigny, France, and on 14-15 October 1998 at Wright-Patterson Air Force Base, Ohio, USA.

The Research and Technology Organization (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote cooperative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective coordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also coordinates RTO's cooperation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of initial cooperation.

The total spectrum of R&T activities is covered by 6 Panels, dealing with:

- SAS Studies, Analysis and Simulation
- SCI Systems Concepts and Integration
- SET Sensors and Electronics Technology
- IST Information Systems Technology
- AVT Applied Vehicle Technology
- HFM Human Factors and Medicine

These Panels are made up of national representatives as well as generally recognised 'world class' scientists. The Panels also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier cooperation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

The content of this publication has been reproduced
directly from material supplied by RTO or the authors.

Printed on recycled paper

Published October 1998

ISBN 92-837-1003-7

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

Alternative Control Technologies: Human Factors Issues

(RTO EN-3)

Executive Summary

Ever since the origins of aviation, the various devices, instruments and aircraft systems involved have always, almost exclusively, been activated by manual controls. At the present time, the high degree of computerisation of all aircraft systems and the generalised use of fly-by-wire means that these systems could easily accommodate non conventional devices such as voice commands, head and eye movement commands etc. All these non conventional devices are often described generically as “alternative control technologies”. These technologies are in fact capable of providing alternative solutions which are also redundant or complementary to manual control in the design of advanced man-machine interfaces. These new technologies could thus contribute to the enhancement of man-machine communications in both military and civil aviation.

The main aim of this Lecture Series is to provide a review of the technologies which can be envisaged at the present time, with their main characteristics, benefits and limitations. These lectures are essentially intended for scientific research workers and engineers involved in the field of man-machine interaction and the design of work stations for aeronautical applications. They may, however, be of interest to others who wish to obtain a summary of recent advances and of the state-of-the-art in this field.

The following questions will be dealt with:

- Operational justification for aeronautical technologies
- Technology and voice command applications
- Technology and head position detection applications
- Technology and eye position detection applications
- Technology and gesture control applications
- Technology and applications of control by biopotentials
- Human factors aspects linked to the integration of these technologies
- Summary and analysis of the benefits obtained

A round table discussion will be held at the end of the Lecture Series.

The material in this publication was assembled to support a Lecture Series under the sponsorship of the Human Factors and Medicine Panel and the Consultant and Exchange Programme of RTO presented on 7-8 October 1998 in Brétigny, France, and on 14-15 October 1998 at Wright Patterson Air Force Base, Ohio, USA.

Techniques de Pilotage Alternatives - Le Facteur Humain

(RTO EN-3)

Synthèse

Depuis l'origine de l'aviation, les différents dispositifs, instruments et systèmes des aéronefs ont toujours presque exclusivement été mis en œuvre au moyen de contrôles manuels. A l'heure actuelle, les systèmes font que l'informatisation poussée de l'ensemble des systèmes avion et la généralisation des commandes de vol électriques pourrait aisément s'accorder des dispositifs non-conventionnels, comme la commande vocale, le mouvement de la tête et du regard, etc. L'ensemble de ces dispositifs non-conventionnels est souvent regroupé sous le vocable de « technologies de contrôles alternatives ». Ces technologies sont effectivement susceptibles d'offrir des solutions alternatives, mais aussi redondantes ou complémentaires au contrôle manuel dans la conception d'interfaces homme-machine avancées. Dans le domaine de l'aviation militaire, mais aussi dans celui de l'aviation commerciale, ces nouvelles technologies pourraient ainsi contribuer à l'amélioration de la communication homme machine.

L'objet principal de ce cycle de conférences est d'apporter une information synthétique sur l'ensemble des technologies qui peuvent actuellement être envisagées, détaillant leurs principales caractéristiques, leurs avantages et limitations. Ces conférences sont essentiellement destinées aux chercheurs scientifiques et ingénieurs travaillant dans le domaine de l'interaction homme - machine et la conception des postes d'équipage en aéronautique. Elles peuvent cependant intéresser d'autre personnes désirant obtenir une synthèse des progrès récents et de l'état de l'art du domaine.

Les sujets qui seront traités lors de ces conférences sont les suivants:

- Justification opérationnelle des technologies en aéronautique
- Technologie et applications de la commande vocale
- Technologie et applications de la détection de position de tête
- Technologie et applications de la détection du regard
- Technologie et applications de la commande gestuelle
- Technologie et applications du contrôle par biopotentiels
- Aspects facteurs humains liés à l'intégration des technologies
- Approche synthétique et analyse des bénéfices attendus.

Une table ronde sera organisée à l'issue de la série de conférences

Les textes contenus dans cette publication ont servi de support au Cycle de conférences 215 présenté sous l'égide de la Commission Facteurs Humains et Médecine dans le cadre du programme des consultants et des échanges de la RTO du 7-8 octobre 1998 à Brétigny, France, du 14 au 15 octobre 1998 à Wright Patterson Air Force Base, Ohio, Etat Unis.

Contents

List of Authors/Speakers

Lecture Series Director: Dr Alain LÉGER
Interface Homme-Machine, Dir. Tech.
Sextant Avionique/MMI
Rue Toussaint Catros, Le Haillan
BP 91
33166 Saint Medard en Jalles Cedex
France

Authors/Lecturers

Dr Timothy R. ANDERSON
Air Force Research Laboratory/HECA
Information Analysis & Exploitation Branch
Wright-Patterson Air Force Base
Ohio 45433-7022
UNITED STATES

Mr Joshua BORAH
Applied Science Laboratories Inc.
175 Middlesex Turnpike
Bedford, MA 01730-1428
UNITED STATES

Dr Grant McMILLAN
Air Force Research Laboratory/HECP
2255 H. Street
Wright-Patterson Air Force Base
Ohio 45433-7022
UNITED STATES

Dr. Graham M. ROOD
System Integration Department, Air Systems Sector
Defence Evaluation and Research Agency
A5 Building Room 2067
Farnborough
Hampshire GU14 0LX
UNITED KINGDOM

Foreword

Currently, manual operation for all kinds of mechanically activated devices designed to control the functions of aircraft systems is used almost exclusively in the aeronautical and space environment, but also more generally in regard to all vehicular control. This has been the rule from the origin of aviation and it is obvious that there are good reasons to explain why this situation has lasted so long.

From the early origins of the species, the superior ability of humans to use their hands in interacting with the environment has been a major characteristic. Actually, mechanical action of the hand on elements of the environment, peculiarly « dumb » ones, such as mineral, and vegetable elements, forms part of the very basic skills of mankind. Quite naturally, the first flying machines were assemblies of wood, fabrics and metal parts. The only « intelligent agent » onboard was the pilot, so it is obvious that the only way to act on the few controls of the aircraft was by mechanical action. Nowadays, even with the introduction of electrical systems and computers, manual control is so robust, efficient and reliable that most interactions with aircraft systems are carried out using the manual mode. Physical contact with the control device provides good and immediate feedback on the action being carried out and generates a high level of confidence in the pilot's mind.

Interacting with other living creatures may, however, proceed from other mechanisms. Animals have many ways to control the behaviour of others without making physical contact, including postures, sounds and facial expressions. Such interaction modalities also exists in humans, but the acquisition of articulated speech introduced a new dimension into the ways of communicating with other individuals and even animals. It should be noted that the semantic contents of words is not the only information provided by speech, prosody and pitch being of great importance as military people recognised a long time ago. The use of voice to control and co-ordinate movements and actions of troops during battles has been the rule from antiquity to modern time. Moreover, heterogeneous redundancy, implying that an identical message transits through different modalities (voice and gesture for instance), is universally used, either to reinforce the content of the message or to complement it. Interestingly enough, the use of such remote control signals requires an « intelligent » agent as receiver.

And there we have the problem. Today, most modern aircraft are totally controlled by computers, which means that some kind of « intelligent agent » is mediating the pilot's actions on the various effector systems. As a matter of fact, the architecture of fly-by-wire aircraft mimics partly, and in a very simplified form, the nervous system of living creatures. All commands sent to the various aircraft systems are electrical signals, thus theoretically suppressing the absolute need for manual control. Some mechanical systems, manually operated, are however usually retained for back-up functions

The computers of the sixties and seventies had limited resources and « intelligence ». Indeed, programmers of this era had a hard time running real time programs with the small amount of memory available on the CPU. They had to put in a lot of effort and imagination to optimise their programs, using assembly languages and « tricks », in order to cope with such limited resources. On the other hand, the human operator is also known to have quite limited resources (perceptual, but also information acquisition, memory access). He is, however, intelligent, and knows how to use various strategies to overcome intrinsic resource limitations.

The situation on the machine side is now completely different. Computers still have poor « intelligence », but they have acquired almost unlimited resources compared to those of the human being. There is now a striking imbalance between a human operator, intelligent, but limited by his resources, and the machine, able to process enormous amounts of data, but still with quite limited « intelligence ». The difficulties encountered at the man/systems interface as a result of this situation have been extensively reported.

In order to improve the communication between the human operator and the machine it appears necessary to work on both sides of the problem: at the man/machine interface and on system design. Most authors agree that working only on the control and display « physical » aspects of the Man-Machine Interface would not produce completely satisfactory solutions. Giving the machine a kind of « Human-Like » intelligence, allowing it to accept high level instructions and to detect the intentions and needs of the operators, is definitely a long term challenge which has been taken up by engineers and cognitive scientists.

Meanwhile, most efforts are spent on the « machine-to-human » relationship, in an attempt to improve information displays and make the information output by the systems easier to perceive and interpret. Of the many concepts of « human-centred » Human-Machine Interface design, the « ecological interface » suggested by Rasmunssen and Vicente some years ago, appears in some aspects to be particularly attractive. This concept states that the interface should be designed in such a way as not to constrain the operator to work at a higher level of control than required by the situation. On the physical side of the interface, this implicitly means that such an « ecological » principle should also be respected with regard to control modalities (and displays). As an example: why should the pilot have to

sequentially designate a series of alphanumerics on a display, when it is far easier to dictate it to an « intelligent » agent (speech recognizer), electronically linked to the aircraft systems?

Introducing « body language » at the interface level is not a new idea. Engineers and scientists have been working for a long time on enabling technology and the way to use it in the aerospace environment. Some of these non-conventional control technologies, as head-trackers or speech recognizers, are starting to be introduced onboard new generation aircraft as the EFA, the Rafale and the JSF.

Quite likely, the major difficulty in integrating more extensively alternative controls into cockpit design will arise paradoxically from the unique adaptive ability of the human being. As a matter of fact, the adaptive nature of the human would probably allow him to perform any task using any control modalities. Also, among individuals, various strategies using various modalities will be developed to successfully perform a similar task. From an engineering point of view, the challenge will be to determine precisely, among the various technologies and combination possibilities what to do, why and how to implement it at the lowest human and economical cost.

To make the best use of these system integration technologies, the ultimate goal should be to allow the user to adopt the most appropriate strategy for him to fulfil his objectives. To remain human-centred rather than technologically driven, great care should be given to identification of the cognitive and sensorimotor « invariants » relative to the use of each technology. On this basis, one of the keys to integrating alternative technology correctly could be seeking to minimise the cognitive and sensorimotor « energy cost » for a given procedure. Trade-off would have to be made between the level of performance required to reach a specific goal and the level of « energy » required to achieve it, including training efforts. Finally, optimising cockpit design by introduction of Alternative Control Technology would mean considering « cost » issues at two levels:

- For the crew, the aim of alternative technology should be to minimise the « cost of control » by making the best use of limited human resources and increasing the global effectiveness of human-machine coupling;
- For the Defence community, the smart integration of these new control technologies should result in training cost reduction, increased operational effectiveness and, eventually, cockpit simplification by using virtual controls.

We can already foresee the limitations of manual controls just looking at the current generation of aircraft under development. Aircraft have used for many years now the HOTAS concept, but the multiplicity of switches, sometimes multifunction, on the stick and throttle raises a lot of questions. Of course pilots can adapt, but this will be paid for through an increase in training needs and higher error rates. Saturation of the very limited and vulnerable short term memory constitutes a major risk here.

It looks as if it is time to increase the resources of the machine in different ways than pure computing power, allowing easier and optimally adapted control of the human over the systems. The motivation is there, the technology is beginning to be mature, operational implementation should now follow shortly.

Dr Alain Leger
Chief Scientist Human Factors
SEXTANT Avionique/ Man Machine Interface
Lecture Series Director

Operational Rationale and Related Issues for Alternative Control Technologies

Dr G.M.Rood

Systems Integration Department, Air Systems Sector

Defence Evaluation & Research Agency, Farnborough

Hampshire GU14 0LX, UK

1. INTRODUCTION

Combat aircraft can, in general, be described as manoeuverable airborne weapons platforms which contain a series of electronic and other systems with which the aircraft is controlled, navigated, weapons selected etc., and a series of systems which provide protection for the aircrew throughout the performance envelope of the aircraft and when emergency escape is unavoidable. Most aircraft platforms have an operational life of over 20 years - some a lot longer - and, in this timescale, although the basic platform does not significantly alter - mainly for cost reasons - the avionics and crew support systems fits can continue to advance a number of generations - which can allow the airframe to retain its operational competitiveness against newer designs

The speed and capacity of future avionic systems, themselves increasing in complexity, will result in the amount of information output heavily increased. This is often all fed to a single pilot who is flying the aircraft close to the ground at around 450 knots or more, perhaps in bad weather at night, and the flying process alone needs continuous monitoring. In addition s/he needs to keep safe control of the aircraft, find the target, select and arm weapons, be aware of, and react to, enemy countermeasures, perform complex operations with smart weapons, etc., all in a degraded environment with high noise levels, high vibration and heat, high 'g' levels, high agility, disorientation, etc. Out of this scenario, one of the primary problems is the amount of data - not necessarily in the right information format for easy digestion - that it is necessary for the pilot to process and the interaction with the displays which s/he will need to ensure that the correct inputs are entered at the right time, and quickly enough, to get the operationally relevant information out.

The more complex the new systems, and this increasing complexity is often needed to counter the increasing subtlety of enemy countermeasures, there is a tendency to need more inputs to a greater number of systems by the pilot and the additional time to carry out these extra operations is not generally available.

The current, and traditional, methods of data input or selection of systems normally require the use of the hands to either switch a system to a particular state or enter data through a key-board. Most current aircraft, both civil and military, make large use of keyboards to enter a wide range of data both on the ground and whilst airborne. Errors do occur in data entry, even under benign conditions, and sometimes can result in serious consequences. In military aircraft, data entry is often an operational requirement in flight and experiments have shown that errors of around 2.2% to 2.9% can occur in high-speed low-level flight [1] and, even in the office environment, typing errors in the region of 1.5% occur, and this is with a full sized keyboard under unstressed conditions and without the need for NBC gloves and the smaller keyboards and key sizes often found in aircraft. Key size differences can occur between a commercial keyboard and a military airborne keyboard - and there are recommended spacings in the Human Factors specification MIL -1472D. Next generation systems may need a larger number of data inputs and to increase the manual input capability of the pilot either requires an increase in 'typing' speed, a larger number of hands or an alternative control technique.

In civil systems errors occur traditionally during high workload periods [2] - often during a runway change required by Air Traffic during approach and, for the military, similar errors could be expected to occur in aircraft which use a combination of military and civil systems in the cockpit (C-130J, E3D, C-17, etc), particularly, perhaps, in the more demanding battlefield support role.

More demanding operations in the current generations of fixed and rotary wing aircraft, particularly at night and in poor weather, have increased the need for more 'eyes-out' operations, which decreases the time for 'head down' or 'head in' viewing time, both for switching operations and for assimilation of information from head down displays. Similarly the speed of operations has led to less time being available for these two operations. Progress has been made towards the assimilation of visual display data through the move towards Helmet Mounted Displays and the time reductions in switching have been achieved through ensuring that the pilot has no need to move his hands from the primary aircraft controls during high workload periods by the use of the **H**ands **O**n **T**hrottle **A**nd **S**tick (**HOTAS**) concept. Using Fitts Law, namely that the time to move the hand to a target (in this case a switch or button) depends only upon the relative precision required, indicates that the movement time - a summed combination of perceptual processing, cognitive processing and motor processing - is in the region of 250 ms (an aircraft moving at 500 knots travels in the region of 80 metres in this time). Thus a time saving of around 250msec is achievable by minimising the hand movements. This generally involves the provision of all of the necessary manual switches on either the throttle top or the control column (stick) top, (HOTAS) or **H**ands **O**n **C**ollective **A**nd **C**yclic (**HOCAC**) - for helicopters - during all critical flight operations. An example of HOTAS controls is shown in Fig. 1-1 for the AFTI F-16 aircraft [3.]

As the capabilities of aircraft will continue to increase through the use of more sophisticated, and a wider range of, sensors, and control through software increases, the ability to control the aircraft systems will inevitably require an even greater number of controls - many of these being necessary, at least in principle, on the HOTAS controls, as many are time critical and need to be operated eyes-out. The rise in the number of avionic systems and the consequent number of manual switching operations necessary during critical phases of operations (eg beyond FEBA and set-up & attack phase of a ground target) has resulted in a gradual increase in the numbers of switches/controls per crew member in the cockpit and this is illustrated in Figure 1-2.

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

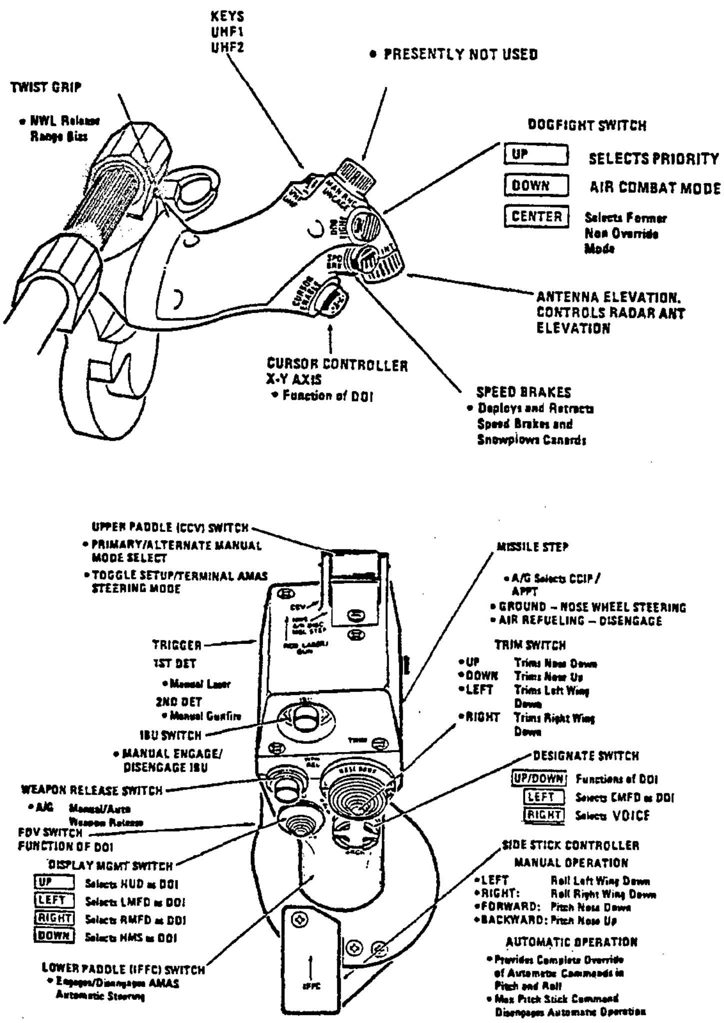


Figure 1-1 HOTAS Controls for AFTI/F-16 Aircraft

The increased numbers of switches and controls results both in longer selection and switching times and with the necessity to look head down into the cockpit to operate the correct switch or series of switches. This has led to the HOTAS concept and, on HOTAS, aircraft of the 1970's design era were using around 16 stick and throttle top functions, and, whilst some aircraft designs in the late 80's still used less than 20 functions, some fixed wing aircraft were up to 33 functions and helicopters up to 40. Figure 1-3 illustrates this trend and Table 1-1 shows the functions allocated to HOTAS for a number of aircraft [3].

There are some indications from aircrew that the numbers of functions are becoming both difficult to remember - needing more training - and sometimes difficult to operate with either standard aircrew gloves or NBC gloves. More complex systems will almost inevitably require more control mechanisms, and the most obvious approach is to increase the number of HOTAS keys - at least for the time critical operations. If the physical space is no longer available on the throttle or stick, the temptation will be to use 'chording' - the simultaneous use of two, or more, (existing) keys to select or operate systems - with an inevitable increase in mental complexity.

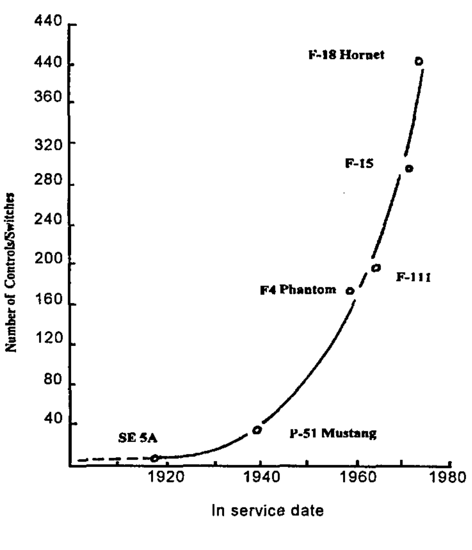


Figure 1-2 Number of controls/switches per crew member

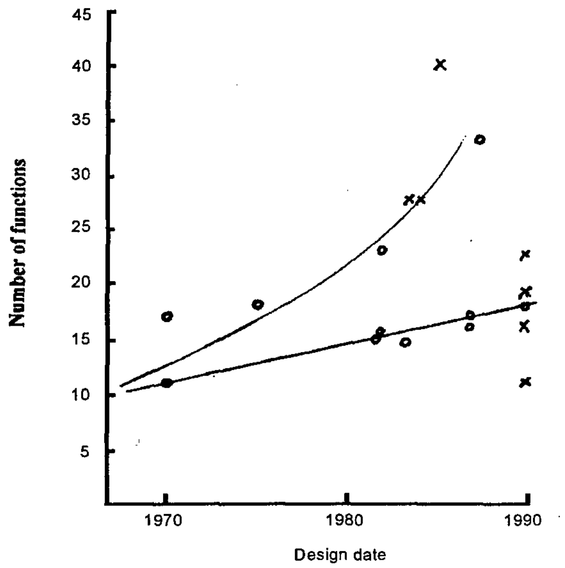


Figure 1-3 Trend of HOTAS Switching

Where the numbers have reached a level where some aircrew are finding some difficulties in remembering the functions of all of the switches, and since it is impracticable to label the switches - it would, in any case, be almost impossible to either clearly read the labels in their position in the aircraft or have the time to read the labels during critical parts of the sortie, - there is no possibility of identifying the correct switch or button if the memory fails or falters for any reason. Since a large percentage of the buttons/switches are for critical aircraft functions, and thus will be time critical, any delay or error can jeopardise the aircraft mission. Further, even if the error is known, the procedures to recover from such errors - if any - inevitably take time. It may not always be clear to a pilot that he has made an error, or that he has pressed the wrong switch or button. If a button is pressed and the expected consequences do not occur, a number of options appear in his mind:

- The switch or button may not have worked:
  - Solution ?? - press again or harder
- The feedback system - if any - may have failed
- The display or function may have failed
- The system may have failed - is there any feedback?
- It may be the wrong button - which one now?

All of these take time, which generally is in critically short supply in these phases of flight. A well implemented **alternative** control input method would provide alleviation of this type of operationally critical problem.

A potential further problem, particularly with the necessary physical positioning of a larger number of switches or buttons is the difference in anthropometric span of the hand & fingers. Not only are there differences in the populations of an individual country, but there are statistical and practical differences between countries - sometimes significant. Currently, a number of countries are accepting female aircrew for combat aircraft, and the differences in HOTAS systems designed for male aircrew may elicit problems for female crew with differing effective digit length and hand-reach anthropometry.

Table 1-1 A sample of functions allocated to HOTAS controls

Aircraft	Design Date	Throttle Functions	Stick Functions	Hand Controller	Total
F15C Eagle	1970	11	6	0	17
F15E-front	1982	9	6	0	15
F15E- rear	1982	0	0	6	6
Tornado IDS - front	1970	4	7	4	15
Tornado IDS - rear	1970	0	0	5	5
F-18 A toD - front	1975	10	8	0	18
F-18 E/F - front	1990	10	8	0	18+
F-18 E/F	1990	0	0	6	6
AV8B+	1989	9	8	0	16
Harrier GR7	1989	17(8)	17(8)	0	16+
Mirage 2000-5	1987	14	9	0	23
Rafale	1988	21	11	0	33
EF2000	1991				
AMX	1982	6	9	0	15
F-16 C/D Falcon	1983	6	8	0	14
AFTI F-16		8	10	0	18
MIG-29		7	12		19
Tiger -rear	1985	14	12		26
- front		14	12		26
AH 64 Longbow-rear	1990	6	13	0	19
-front		0	0	11	11
EH101	1984	19 (14)	21 (12)	0	40
RAH66 Comanche	1990	14	8	0	22
MV22 Osprey	1988	9	7	0	16
A330 Airbus	1990	0	3	0	

Table 1-2 shows an example of the differences in hand length of a number of countries and of a number of trials. The average hand length for males is 191.65 mm with an average spread of 48 mm. Standard deviations are in the region of 9 mm, which, as an estimate, would allow a HOTAS mounted set of switches and buttons to be designed to be used by perhaps some 70% (>1 sd) of the pilot population without undue difficulty. The remaining 30% may need to make some sliding movements around the stick or throttle to accommodate the full range. The female average hand length, however, is an average of 176.3 mm with a spread of 42.5 mm and an sd of 8.6 mm. The difference in mean length is some 16 mm, which could provide some difficulty in design of HOTAS controls which must be operated by both genders.

Table 1-3 supports this hypothesis with figures comparing, in more detail, differences between UK male and female hand dimensions [4]. As an indication of the potential problems, the distance from the 'hand crease' - representing, in this case, the apex of the HOTAS grip - to the finger tips displays an average difference of 1.2 cm. If a wider range of male and female crews need to be accommodated, then this difference may be increased to over 3 to 4 cm. Similarly for span between the thumb and the individual digits, which gives an indication of the ability to operate a thumb switch and another with one of the other digits average differences of around 1.3 cm are apparent.

Table 1-2 Hand Length Data

	Date	Sample	mm	sd	Range	Spread
Male						
UK Military	1982	300	191.30	9.71	169-224	55
Canadian Military	1974	565	191.90	8.78	170-212	42
German Air Force	1966	1006	189.10	8.70	168-210	40
British Army	1970-75	2000	193.00	10.30	159-219	60
US Army	1970	1482	192.00	8.70	172-214	42
US Army	1966	6682	190.30	9.60	169-214	45
US Air Force	1970	148	197.20	9.30	173-228	55
French Army	1973	793	189.00	9.00	174-205	5th-95th% range
UK Civilian	1981	300	191.00	8.30	165-219	54
mean values			(191.65)	8.27	(159-228)	48
Female						
UK Military	1982	187	176.10	8.07	159-197	38
US Army	1977	1331	174.40	9.00	155-196	
US Air Force	1970	211	179.30	8.60	157-205	43
UK Civilian	1980	92	177.50	10.10	161-194	5th-95th% range
UK Civilian	1981	200	174.20	7.20	152-195	43
mean values			(176.30)	(8.6)		42.5

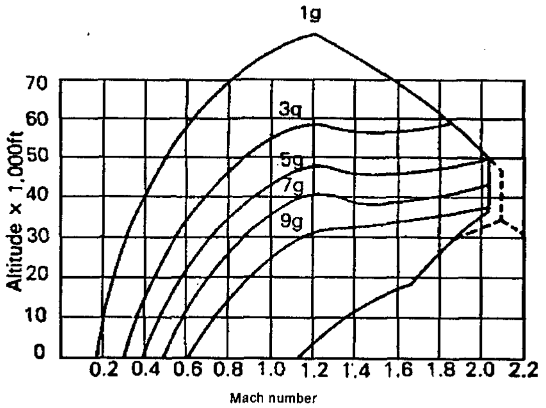
Table 1-3 Details of Hand Dimensions

		Male			Female		
		Mean	Sd	Range	Mean	Sd	Range
Finger number to hand crease							
Digit 2	Left	12.26	0.81	10.2-14.4	11.16	0.70	9.4-13.4
	Right	12.21	0.79	10.2-14.9	11.17	0.69	9.3-13.1
Digit 3	Left	13.51	0.92	11.1-16.0	12.12	0.79	10.3-14.4
	Right	13.40	0.87	11.3-16.5	12.09	0.78	10.3-14.6
Digit 4	Left	12.44	0.95	9.8-15.0	11.00	0.86	9.0-13.5
	Right	12.31	0.90	10.0-14.9	10.97	0.85	9.2-13.6
Thumb	Left	6.02	0.50	4.7-7.6	5.50	0.44	4.3-6.9
	Right	6.11	0.48	4.9-7.8	5.62	0.43	4.2-7.0
Digit 5	Left	9.13	0.98	7.0-12.3	8.28	0.80	6.2-10.9
	Right	9.49	0.90	7.1-11.8	8.31	0.89	6.3-11.1

Similarly the true digit lengths - the length of each finger -is shorter for females by around 5 mm and the curved hand length is shorter in females by some 1.65 cm. Many of these differences may be able to be accommodated by good design, but there must be a high probability that, in current designs, and in future designs where the increasing number of controls surfaces will perhaps result in physically smaller switches and buttons, the potential competition between switch numbers and available surface area, as numbers of switches or tactile controls compete with surface area, will play a more significant limitation.

2. HEAD POINTING

Currently, the majority of aircraft carrying out a missile attack on a ground or airborne target must point the nose of the aircraft towards the target in order to suitably align the enemy aircraft on the weapon aiming displays on the HUD to lock-on the weapon prior to firing. This is not only a time consuming approach, but may require the aircraft to perform tortuous manoeuvres in pursuit of the also manoeuvering target aircraft. Figure 2-1 illustrates the sustained and instantaneous manoeuvre capability that is currently required from an air-to-air combat fighter, in this case the F16.



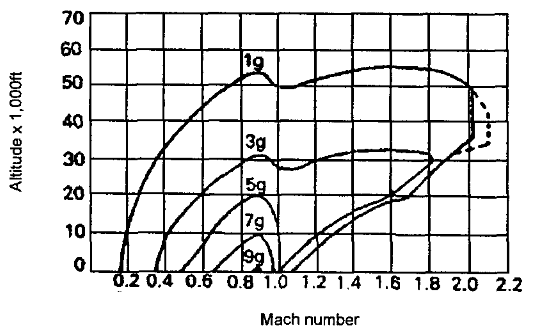


Figure 2-1 Instantaneous (upper) and sustained (lower) manoeuvre capability of the F-16/79

Unfortunately the human body, being developed over a few million years for a less stressful environment, does not respond well to these violent manoeuvres and technologically complex and ingenious methods of protecting the body must be employed. Currently airframe soft limits in the region of 9 'g' are in use in current production and future aircraft and the protection of the crew to these levels is complex and cumbersome.

The emergence of the technology, over the last 15 years, to allow flight worthy Helmet Mounted Displays (HMD) [5, 6, 7] and the development of accurate flight worthy Head Pointing Tracker Systems (HPS) has allowed methods other than manually boresighting the aircraft, to be used to enhance weapon delivery techniques.

Future-current and next generation weapon systems, particularly air-to-air close combat engagements, will be able use an alternative form of control system that will integrate the HMD, the HPS and the missile seeker head, Figure 2-2.

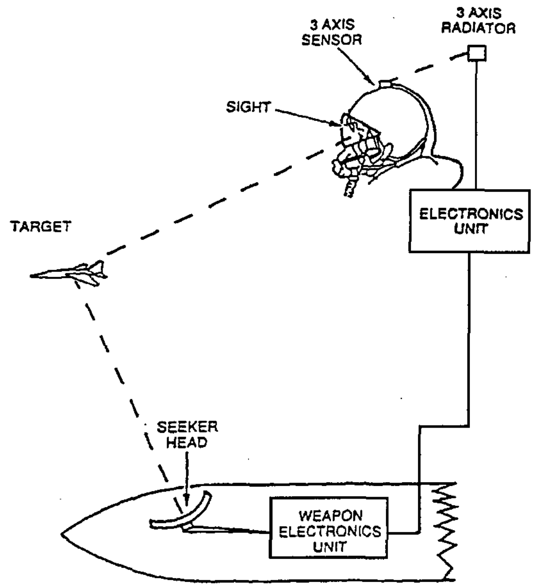


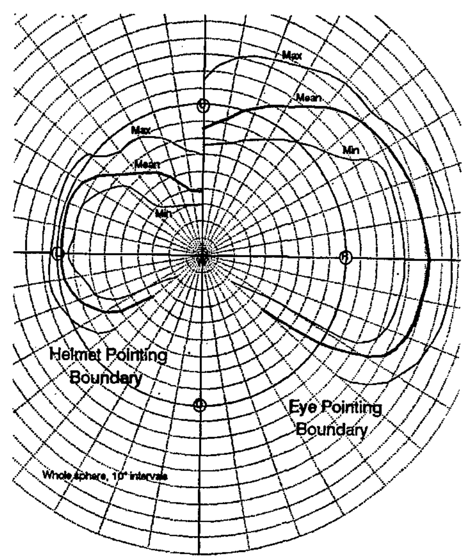
Figure 2-2 Use of head pointing in an off-bore-sight capability weapon system

This will enable the missile seeker to be driven by the head pointing system to look in the direction that the pilot's head is pointing, and, as the pilot sights the target aircraft in his helmet mounted sight, for the missile to lock-on and be fired at high off-boresight angles, without the necessity for violent manoeuvering of the aircraft. Flight trials both in the USA, where live missiles have been fired at drones (BOXOFFICE) and in the UK, where air-to-air close combats have been carried out in 1 v 1 trials (JOBTAC) significant reductions in target acquisition and engagement times are apparent.

The use of a helmet mounted display and head tracking system in an F16, combined with a missile capable of acquiring targets of over 60 degrees off-boresight, has allowed, in live firings against a QF 106 target drone at 0.7M, successful intercepts at 57 degrees off-boresight whilst the target was manoeuvering at 5g. Similarly, in one-on-one or two-on-two air-to-air combat between a MIG-29's fitted with a simple Russian helmet mounted sight and using a AA-11 (Archer) missile, and F16's with no helmet sight, the MIG-29 was able to attain the major number of first shot missile releases by use of the Helmet sight system. To pass the head position information to the missile seeker, the MIG-29 used an electro-optical head tracking system. [8].

Similarly, at Farnborough in the UK, trials have been flown of one-on-one combat in a Jaguar, using a captive AIM9L and a standard Mk4 UK flying helmet fitted with a simple DERA/GEC sight providing weapon systems information through an LED display and an AC electro-magnetic head tracker. Target acquisition and engagement times were significantly reduced, with off-boresight acquisitions up to 60 degrees being achieved.

As with most systems, however, whilst there may be significant operational shorter term advantages, there are also some longer term restrictions in the systems use of Helmet Mounted Head Pointing Systems. One of those comes from the inability of a correctly strapped-in pilot to move his head much more than 90 degrees to the left or right. Figure 2-3 shows the head pointing envelope of a pilot, in full flying clothing, in a fast-jet strike aircraft cockpit, and, whilst the envelope is acceptable, it is limited by the available head movement of the human body. If, however, a further alternative control method, in the form of eye-tracking is utilised, then the useable envelope is significantly increased. This will allow, on average, tracking to around ± 140 degrees in the horizontal plane, compared to ± 90 degrees for head tracking and up to 90^0 in the vertical planes, compared to 55^0 with head tracking (in an aircraft with the restricted rearward and upward movement of the head from an Ejection Seat Headbox).



Comparison between helmet pointing and eye pointing envelopes

9 subjects
Mk10 ejection seat, tight harness
Tornado cockpit
Mk4 helmet, summer AEA and LSJ
(from RAE Tech Memo FS(F)-680 1987 by T K Manning)

Figure 2-3 Head and Eye tracking Envelopes

Thus, it should be technologically possible to targets in the rear hemisphere - or, at least be able to input information into the weapon systems as to the position of target aircraft outside of the conventional radar systems field-of-regard or missile seekers FOR [unless missile design changes] - but not, perhaps, outside of next generation thermal sensors FOR. The Russian Vympel Design Bureau is reported as having tested a rear engagement capability in 1993 on a Sukhoi Su-27. The control authority of thrust vectoring allowing a rearward shot without the missile losing control as it initially flies backwards, [8].

Head Tracking can also be used to designate ground targets from the air, or to point narrow FOV sensor systems at targets - and these generally replace manual control systems that are displayed on a HDD. Hunting for a target, in a moving aeroplane, with a narrow FOV sensor (likened to looking for a target through a straw) can be difficult in the best of conditions and may take longer than is acceptable. By the use of either a Helmet Sight with Head Tracking, or with the addition of Eye Tracking, this type of operationally essential process can be considerably shortened and higher accuracies attained. UK trials have linked together such a system enabling the FLIR sensor in TIALD (Thermal Imager and Laser Designator) to be located directly on

a target of opportunity using a helmet sighted system in conjunction with the head tracker.

3. EYE TRACKING

Eye Tracking has also some similar potential within the conventional cockpit or cabin, particularly with the use of large picture displays. These displays can either be in use in rotary or fixed wing strike aircraft, or in surveillance or Command & Control type aircraft. The problem lies in the use of a cursor in a large, and often cluttered, display, where the position of the cursor on the screen is not always immediately clear. For small FOV displays (say 20 deg x 20 deg) the cursor position can be determined more easily as it lies generally within the foveal cone of the eye and conventional manually controlled mice or joy-sticks are adequate. In a larger display, however, it can need considerably more scanning to find the cursor prior to repositioning it - with the obvious time delays. With conventional cursor control, it is necessary to find the existing position of the cursor in order to know which way to move the manual control to reposition the cursor at its new point By the use of eye tracking, however, it will be possible to reposition the cursor by the combination of fixing the eye on the required point and commanding the reposition with either a manual control or by the use of a voice command. This could also be used to reposition target boxes or similar designators in large screen displays, and combinations of eye tracking for coarse control and manual for fine control are feasible options. This combination of eye designation, manual fine control and target box labelling by voice command has the potential to provide significant reductions in aircrew workload.

4. VOICE CONTROL

Voice control or Direct Voice Input (DVI) has a large potential for Alternative Control Techniques. In the HOTAS case, the problems may lie in the inability to remember either the position of the switch or the name of the function to be operated - more probably the former than the latter. With the use of voice command to switch the system, the problem of memorising the switch or button positions is effectively nullified, and only the lesser problem of remembering the functions is left - in practice this should significantly reduce errors. Again, in practice, as with most **alternative** control technologies, it would be wise to retain redundancy in the system and allow operation by either manual and/or voice operated controls - pilot preference being allowed depending upon sortie patterns and phases. By using both systems, the number of manual operations on the HOTAS controls could be significantly reduced and HOTAS used for the time critical functions only, rather than its current potential for over-use - as there are no alternative control techniques to replace manual switching.

The use of voice control or Direct Voice Input (DVI) to select and switch systems has been discussed for a number of applications and is probably the lowest risk of alternative control technologies. One major advantage over manual hard or soft key control is in being able to enter a, sometimes complex, hierachical control structure at any point. In most current systems (navigation, attack, TV-TABS, etc.) it is necessary to page through the levels of a hierarchical menu to reach the level required. In the RAE (now DERA) Tornado flight trials, DVI was used on the navigators TV-TABS and it was possible to access different levels of the navigation hierarchy directly with potential time savings. Whilst later systems have a less time consuming approach to the ability to access deeper parts of the system hierarchy, there remain structural problems with this approach, and whilst considerable ingenuity has been expended on reducing the number of button presses to access the required information, only manual keyboarding or voice control will allow direct access to the functions

Other areas that would benefit from the use of DVI are in the areas of Radio Channel selection. Currently, when a pilot needs to talk to a new controller, ground control, approach, tower, FAC, etc, it is necessary to obtain the frequency and select it on the appropriate radio - VHF/UHF/HF etc., before transmission. This process of obtaining the required controller, say Paris Orly approach, leads through mentally remembering the required frequency or looking up the frequency, through manual selection of the frequency on the appropriate radio and finally transmitting and talking to Orly approach - the person you first thought of - is unnecessarily time consuming - and in many military operations time will matter. Voice command will shorten this process by asking, in a single operation, for Paris Orly Approach directly - the avionics will do the rest by recognising the request and having the frequencies already allocated to the controller in the avionics. In military operations, particularly during helicopter attack operations it is not

5. UNMANNED AIR VEHICLES (UAVS)

Over the next decade there is likely to be an increasing transition from air based cockpits to ground based cockpits for use with man-in-the-loop Unmanned Air Vehicles (UAVs). In the manned aircraft, the trend is likely to be, at least in a large number of air-to-ground operations, to isolate the human crew, as much as possible from the risks associated with combat areas. The natural trend, which is already visible from recent conflicts, is to produce stand-off weapons, either autonomous or with a man-in-the-loop control capability. Currently this is done from an airborne platform situated far enough from the target to minimise the risk of loss of, or damage to, the aircraft. As data links improve, by increased distance, immunity to jamming and increased bandwidths, the controlling site will be able to move to larger aircraft platforms and finally to ground borne stations. In each of these ground stations (ground or air based), control can be of either UAVs which are intended to fly returnable missions - or UAVs which are not intended to return to base.

Movement of the control station to the technically, and environmentally, more friendly ground station has a number of obvious advantages. Noise, vibration, heat and those discomforts and partial disablers associated with aircraft manoeuvres - high 'g' for example - are not present and the encumbrancies necessary for aircrew protection -laser protection, flying helmet, oxygen mask, 'g' suit, NBC personal equipment etc., - are eliminated. Other factors, such as displays equipment, do not require the airborne equipments limitations on mass & volume to be implemented, nor do associated issues such as display brightness and display power. This should allow Commercial Off the Shelf (COTS) avionics equipment to be more utilised which will significantly support the affordability of these type of military operations.

Consequently, the use of Alternative Control Technologies to supplement the natural human performance, often in terms of speed and accuracy, rather than compensate for the inadequacies and compromises that are essential in the cockpit environment, are more viable.

For instance, head-tracking systems are not exposed to unwanted motion from ground induced turbulence during ground attack sorties, voice system recognition rates improve in a low noise and vibration free environment, eye tracking devices will not

require the complex integration into the airborne flying helmet and devices that are sensitive to environmental infra-red emissions (eg sunlight) can be more readily used - if appropriate.

The benefits of using alternative control technologies are not only apparent in the severe military air environment. The ability to operate more naturally with avionic and military systems, even in the more benign environments of the surveillance aircraft or the ground-borne UMA/UAV cockpit, should provide significant benefits to military operations.

6. CONCLUSIONS

Future manned cockpits will inevitably have more complex avionic fits to cope with more demanding operational scenarios and aircraft roles, and there will need to be an advance in the way that aircrew interface with the aircraft systems in order to enable efficient control between man and the rising complexity of aircraft systems. The number of manual control systems, including buttons, keyboards, and switches, is reaching a point where training aircrew to remember the phases and modes of switching could become both a significant proportion of operational training cost and also have flight safety implications. Similarly the increasing number of switches on HOTAS controls has the potential to heighten confusion rather than provide solutions. What is required are **alternative** methods of inputting data to aircraft avionic systems, particularly if they provide a more natural, and quicker, interface. A simple example of this is in the use of voice input as an alternative to remembering and dialling up radio frequencies. A single command phrase - Farnborough Tower - for instance, replaces, essentially, a three segment approach - remember frequency, dial frequency and call controller on that frequency. Of the more mature alternative control technologies, voice recognition and head tracking are both in operational flight and experimental flight - depending upon the level of sophistication of the technology - and are both technically mature enough for full operational use, with research on the next generation, higher capability, systems in progress.

Eye based control is laboratory mature, and used for assessing eye movement in simulators, and, with development, has the potential to integrate effectively in the operational environment with head and voice based control. Gesture and biopotential are probably the least mature, but provide potential for the longer term aircraft systems (2020) and may be particularly of use in ground based cockpits of man-in-the-loop UAVs.

All systems in a civil and military aircraft must provide some tangible operational benefit - particularly in retrofit cases - and both head and voice based control are expected to provide that benefit in the third generation aircraft (Eurofighter and Rafale). This would be supplemented, in due course, with eye based control, particularly in the air-to-air engagement role, but, also, to a lesser extent, in the air-to-ground role.

The benefits of alternative control techniques lie in a more natural interface with the aircraft, improved speed of operation and reduction in training overheads.

Released from the constraint of only one communication channel with the aircraft systems - manual - the use of alternative control technology invites aircrew, aircraft and systems designers, and others, to be more imaginative in their interaction with the aircraft and systems, using these alternative controls as appropriate to the operational benefits and needs. Such alternatives are not intended primarily to replace manual controls but to supplement manual systems and to provide alternatives, to be used as the occasion requires. Aircraft systems, however, need to be practical, to retain as simple an interface as the technological complexity of the systems allows and be operated by aircrew with a wide range of capabilities. This should ensure that the use of these alternative controls is balanced by the aircraft designers natural, and often historically justified, inherent scepticism of the useability of new technologies.

REFERENCES

1. White G. and Becket, P. "Increased Aircraft Survivability using Direct Voice Input AGARD CP 1983
2. FAA Human Factors Team Report on: The Interfaces between Flight Crews and Modern Flight Deck Systems June 1996
3. Flight Vehicle Integration Panel Working Group 21 Glass Cockpit Operational Effectiveness AGARD-AR-349, 1996
4. Gooderson C.Y.*et al* "The Hand Anthropometry of Male and Female Military Personnel" Army Personnel Research Establishment Memorandum 82M510, 1982
5. The Man Machine Interface in Tactical Aircraft Design and Combat Automation AGARD Conference Proceedings No 425 (AGARD-CP-425), 1987
6. Combat Automation for Airborne Weapon Systems: Man Machine Interface Trends and Technologies AGARD Conference Proceedings 520 (AGARD-CP-520), 1993
7. Advanced Aircraft Interfaces: The Machine Side of the Man Machine Interface AGARD Conference Proceedings 521 (AGARD-CP-521), 1992
8. Aviation Week 16 October 1995

THE TECHNOLOGY OF SPEECH-BASED CONTROL

Timothy R. Anderson
Technical Advisor
Information Analysis and Exploitation Branch
Crew System Interface Division
Air Force Research Laboratory
Wright-Patterson AFB, OH 45433-7022

1. ABSTRACT

This lecture will present an overview of current speech recognition/control technology being utilized for aerospace applications. Common approaches in the areas of signal acquisition, signal processing, and pattern matching will be presented. Pattern matching algorithms for speech recognition/control can be characterized as pattern recognition approaches and acoustic phonetic approaches. The most common pattern recognition approaches used today are the hidden Markov model (HMM) and neural network. The strengths and weakness of the various approaches will be examined.

2. INTENTION OF THE TECHNOLOGY

Current speech-based control systems are the maturest of those alternative controls discussed in this lecture series. Although research in this area goes back over 25 years [1], applications are only recently becoming widespread and accepted by the user community. This is due for the most part because of both limits in the technology and the very high expectations of the technology. It must be highly accurate, robust, and reliable to meet user needs and expectations. Speech-based control systems must be easy to use, that is, transparent to the user. The system should adapt to the user; not force the user to adapt to the system. In the following sections, a brief tutorial of terminology and components of speech-based control will be presented.

When discussing automatic speech recognition (ASR) systems, it is convenient to subdivide them into classes according to the problems they address. Systems are usually first divided according to the number of speakers they recognize.

Speaker-dependent systems can recognize speech from only one speaker, the speaker that trained the system. Speaker-independent systems recognize speech from many speakers, not only the speaker that trained the system.

The next subdivision that occurs for ASR systems is based on how they handle word boundaries. Isolated word recognition systems require a 100-250 ms or longer pause inserted between spoken words. Connected word recognition systems require a very short pause between words. Continuous speech recognition systems require no pause between words and accept fluent speech.

An additional subdivision that occurs for ASR systems is based on the size of the vocabulary or number of words that the system can recognize. Vocabulary size is usually divided into small (less than 200 words), large (1000 to 5000 words), very large (5000 words or greater) and unlimited (greater than 64000 words).

When defining a vocabulary for a specific task, a grammar may be developed that specifies which words may follow other words. This syntax, when incorporated into the recognition algorithm, has the effect of reducing the total number of words that must be considered by the recognizer at any one time. This improves both the speed and accuracy of the recognizer. Perplexity is a common metric used to determine the complexity of a grammar. Perplexity is defined as the average branching factor of the grammar or, stated another way, the average number of words that can follow each word in the grammar. The larger the perplexity of a grammar, the more difficult the recognition task.

Which combination of characteristics is best? The answer depends on the particular application that one is trying to accomplish with speech-based control and the characteristics of the user, task, and environment.

3. OVERVIEW OF APPROACHES

Speech generation is described by means of the "Source-Filter" model: a source of sound energy, which may be regular pulses from the vocal chords, or random fluctuations in the pressure of air being forced though a narrow constriction, is applied to a cavity with many resonant frequencies (i.e. the vocal tract). The frequencies and bandwidths of the resonances are determined primarily by the shape of the tongue, but also to some extent by the jaw position, lips and velum.

In normal usage, speech carries several different kinds of information. As well as the semantic content, there is also information about the physical and emotional state of the speaker and cues to control the dialogue between speakers. The microphone and subsequent signal conditioning modify the speech signal. In control applications of speech recognition, only the semantic content of the speech signal is required, so all the other kinds of information tend to act as perturbations that reduce the recognition performance. The speech signal could also be used to monitor the speaker's physical or emotional state (see "Applications of speech-based control", this volume).

Automatic speech recognition can be viewed as a pattern recognition task that maps an input speech waveform to its corresponding text. Although a wide variety of specific components and processes have been used, all speech recognition systems consist of combinations of the following functional elements:

- Signal acquisition -- microphones of various styles and frequency responses.
- Signal processing -- digital signal processing algorithms that identify or quantify the speech signal.
- Pattern Matching -- algorithms that transform the processed speech into a text string of the recognized speech.

Each of the components will be described in the following sections.

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

3.1. SIGNAL ACQUISITION

The speech signal is characterised by variation of energy with both time and frequency. The frequencies of interest lie between about 100 Hz and 8 kHz, although a narrower bandwidth can suffice to carry intelligible speech. Ordinary telephones transmit frequencies from 300 Hz to 3400 Hz. In the time domain, the most rapid variations typically occur over durations of a few milliseconds. At the upper end, some vowel sounds, and other features, may remain relatively stable for 100-200 ms.

The most commonly used microphones are the close-talking headset microphone and the telephone handset, although other possibilities are lavaliere, desktop, and array microphones. Each microphone presents it own unique challenge because of the various frequency characteristics and signal strengths of the microphone or the mode in which it is used. For example, a desktop or array microphone allows the user to walk around the room, resulting in various signal strengths as a function of the user's relative position to the microphone. These challenges are even greater for speaker-independent systems where different microphones were used for training than those used in the desired application.

In military aircraft cockpit applications, the microphone is included in an oxygen mask. The transfer function is then due to the influence of the microphone and the acoustic cavity. The resulting transfer function is widely imperfect and, even if it is sufficient for speech communications, it must be balanced (flat) for speech recognition. One way to solve the problem is to incorporate pre-emphasis filtering in the signal parameterization chain. The second solution is to use microphones of better quality and to design new oxygen masks, in order to provide a transfer function as flat as possible. This second solution is obviously more complex than the first one, and could be adverse to some constraints the oxygen mask must respect. For example, under over-pressure, the pilot's security and integrity remain more important than speech recognition. In the case of rotary wing applications, the same problem occurs as soon as oxygen masks are used; but in some rotary wing applications, the pilot uses a differential close-talking headset microphone. Due to the environmental noise, a pilot puts the microphone as close as possible to his mouth. In this case, the acquired signal involves electronic saturation. Such a problem can be easily solved by training in the same conditions (without noise but with a microphone position analogous to that in real flight), or by adjusting the audio return so that the pilot positions the microphone further from his mouth.

Gain control is also a practical problem, which can greatly effect speech recognition performance. Analog tools provide speech acquisition, but in order to compute speech features to be recognized, an analog-to-digital converter is required. This analog-to-digital converter involves a processing gain that must be adjusted in order to avoid overflow during numerical computations. But since speech is a highly varying signal, gain adjustment must be accurate. If the gain adjustment is not dynamic, some speech sounds will be coded over a very few bits, without using the dynamic range of the converter and introducing quantization noise. In order to optimize the quantization dynamic range, an automatic and adaptive gain control is required. One would think that classical Automatic Gain Control (AGC) methods are sufficient, but this is not the case: if the speech level is too variable, the AGC can be adverse to speech recognition.

In most systems analog-to-digital conversion is performed at sampling rates of 8000 Hz or higher. The speech power in specific frequency bands is then estimated with Fast Fourier Transforms, digital band-pass filters, or some auditory modeling techniques. These signal-processing techniques will be discussed in the next section.

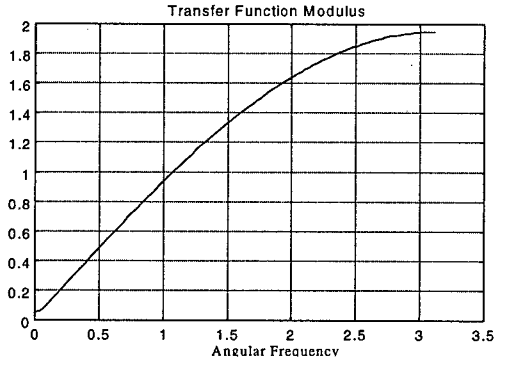


Figure 1 Pre-emphasis filter transfer function

3.2. SIGNAL PROCESSING

Before the pattern matching stage of speech recognition can take place, it is necessary to transform the speech waveform into a more tractable representation. This is necessary to reduce the quantity of data that the pattern matcher must handle. A second, but related, purpose is to extract those features of the speech signal which carry the information that discriminates between words, while eliminating features that carry other types of information. Information relating to the pitch of the signal is generally discarded for purposes of speech recognition (at least in European languages - pitch may be important in tonal languages such as Mandarin).

In most cases, a discrete pre-emphasis filter that compensates for the natural decrease of 6dB/octave due to human speech production precedes digital speech processing. A classical filter is given by the following formula:

$$H(z) = 1 - 0.95z^{-1}$$

whose transfer function is shown in Figure 1.

Although there are many different ways of representing the speech signal, most of them have certain features in common. Almost all techniques produce some kind of representation of the short-term power spectrum over a period of 5-30 ms.

Speech is a quasi-stationary signal; the spectrum may be approximately constant over periods of a few tens of milliseconds. It may also change rapidly within a few milliseconds, in plosive consonants, for instance. The purpose of windowing is to select a finite portion of the signal, which may be considered stationary, for analysis. The length of the window must be a compromise between spectral and temporal resolution. A long window will give a high-resolution spectrum, but may hide the more rapid changes in the signal, whereas a short window will reveal the temporal structure more precisely, but blur the spectral characteristics. Window lengths of between 10 and 30 ms are commonly used for speech analysis.

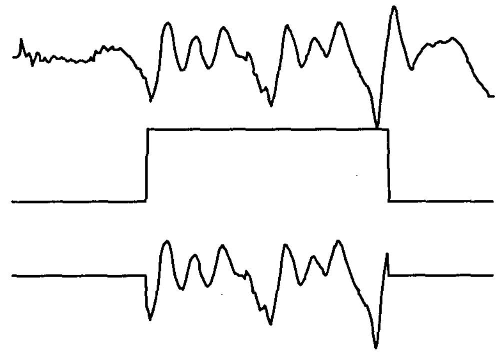


Figure 2 Rectangular window: Original signal (top); rectangular window (middle); windowed signal (bottom).

Mathematically, windowing is equivalent to multiplying the signal by a function that has a value between 0 and 1 within the window and 0 at all other times. The simplest window is the uniform, or rectangular, window of length N samples:

$$\begin{aligned} w(n) &= 1, \quad n = 0,1,\ldots.N-1 \\ &= 0, \quad \textit{all other } n \end{aligned}$$

Figure 2 shows a frame of a signal extracted with a rectangular window. The temporal properties of the signal have been changed by this process, i. e. the new signal is zero outside the window. As a consequence, the spectrum of the signal is also inevitably changed. The spectrum of the windowed signal is obtained by convolving the spectrum of the original signal (assumed stationary) with the spectrum of the window [2]. The window spectrum is similar to a low-pass spectrum, with a broad main lobe at low frequencies and attenuated side-lobes at higher frequencies. The ideal window response will have a very narrow main lobe and large attenuation in the side-lobes. This can only be achieved by using a very long window, which defeats the object of using a window in the first place.

The rectangular window has a narrow main lobe for its length, but the attenuation in the side-lobes is very poor, only around 20 dB. A broad main lobe can be tolerated more easily than poor side-lobe attenuation, as the former reduces the local resolution while the latter spreads energy from distant parts of the spectrum. (In speech signals adjacent frequencies tend to be quite highly correlated anyway, so the local resolution is less important.) For this reason, many attempts have been made to design windows that reduce the side-lobes as much as possible. This is achieved by tapering the edges of the window in some way. Figure 3 shows the widely used Hamming window, which is described by the following:

$$\begin{aligned} w(n) &= 0.54 - 0.46\cos\left(2\pi n/(N-1)\right), \quad n = 0,1,\ldots N-1 \\ &= 0, \quad \textit{all other } n \end{aligned}$$

The side-lobe attenuation of the Hamming window is about 30 dB greater than that of the rectangular window. Note, however, that the samples towards the edges of the window are considerably attenuated, so it is important to overlap the windows for successive frames. If this were not done, important features of the signal that happened to fall on the boundaries between frames would not be given due prominence in the final signal representation.

Many other window designs are possible, although only a few are commonly used, such as von Hann, Hamming, Kaiser, and Blackman [3]. Which is best is dependent on the application, though the Hamming window is probably the most common in speech recognition front ends.

Following windowing, the frame is analysed by one of many possible methods, resulting in a string of about 10-20 numbers called a vector. In many cases, elements derived from the rates of change of the basic vector elements are added to the vector. The following paragraphs describe the most commonly used signal representations and discuss their various advantages and disadvantages.

The simplest representation of the speech signal is achieved by passing it through a parallel bank of band-pass filters (see Figure 4). There are usually between 10 and 20 filters, covering the band from 200 Hz to 4 kHz. The bandwidth of each filter varies according to its center frequency, typically from 200 Hz at the low frequency end to 500 Hz at the high frequency end. The output of each filter is rectified and smoothed with a low-pass filter (cut-off usually about 25 Hz). The resulting value is sampled at the frame rate (50-100 Hz) and may be used directly or (more usually) compressed by taking its logarithm. An equivalent representation can be achieved by means of a Fourier transform followed by summation of the components within each frequency band.

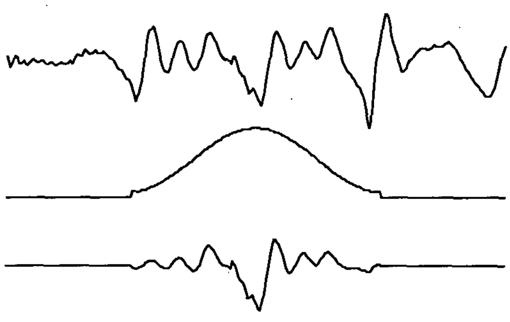


Figure 3 Hamming Window: Original signal (top); Hamming window (middle); windowed signal (bottom)

An alternative way of representing the spectrum is to derive the parameters of an all-pole filter having the same response as the vocal tract at that point in time. This representation is known as Linear Prediction because of the technique used to calculate the filter coefficients using a linear combination of past waveform samples to predict the next sample. Many different methods exist to calculate these filter coefficients. See Rabiner and Juang [4] for a review of these different techniques and their advantages and disadvantages.

Several signal representations model, with varying degrees of accuracy, the processes believed to be used by the human auditory system. The motivation for this derives from the fact that speech has evolved in conjunction with hearing and therefore, the nature of speech is heavily dependent on the capabilities of the ear.

Perceptual Linear Prediction (PLP) [5] implements three concepts from hearing to estimate the auditory spectrum: (1)

the critical-band spectral resolution, (2) the equal-loudness curve, and (3) the intensity-loudness power law. The auditory spectrum is then approximated by an all-pole model (the same basic idea as Linear Prediction discussed above).

The filter bank described above may be regarded as a very low-resolution auditory model. The main analogies with the human ear are that the bandwidth of the filters increases with frequency (the mel scale, [2]) and the amplitude response is logarithmic. At the other extreme, a full auditory model may have 100 channels and provide an output that mimics the firing of the nerves that carry signals from the ear to the brain. The computational power required for this kind of signal representation is very high.

The so-called cepstrum is derived by transforming the speech signal into the frequency domain with a Fourier transform, taking the logarithm of the power spectrum, and then using the inverse Fourier transform to return to the time domain. This gives a representation akin to a spectrum, but the horizontal axis is time (hence the name "cepstrum"). It is easy to distinguish between the pitch component and those components that represent the shape of the vocal tract.

Several of the basic signal representations may be greatly improved by subsequent processing using a technique known as Linear Discriminant Analysis. This is an optimization technique, applied during the development of the recognizer, or possibly during the training of the word models, which is used to find the combinations of channels and/or channel deltas which are best able to discriminate between the words of the vocabulary. The best known version of this technique is the IMELDA (Integrated mel scaled Linear Discriminant Analysis) transform [6]. The effect of applying the transform is to concentrate information into a small number of channels with little correlation between channels.

While a general transform can be derived for a given recognizer, this technique can be optimized for specialised applications, such as military aircraft. This gives a worthwhile improvement in performance.

The analysis of the human cochlea takes place on a nonlinear frequency scale, known as the Bark or mel scale. This scale is linear to about 1000 Hz and is approximately logarithmic above 1000 Hz. It is common to perform such a frequency warping for representations of speech. The most commonly used method of feature representation is that of mel-frequency cepstral coefficients or MFCCs [7]. MFCCs are generally computed every 10 ms by first performing a spectral analysis using a Fast Fourier Transformation on a window of 20 ms of speech. The spectrum is then warped using the above-mentioned mel-frequency warping. The logarithm of this warped spectrum is taken and followed by an inverse Fourier transform. The result is called the mel-cepstrum. By keeping the first dozen coefficients of the cepstrum, the spectral envelope information is preserved. The resulting features are the MFCCs.

The Fourier transform is one of the basic signal analysis tools relevant to analyzing stationary signals. But in the case of short-duration phenomena such as unvoiced plosives (/p/, /t/, /k/), the Fourier transform becomes less accurate. The wavelet transform, which appeared in the last decade, has been introduced in order to process such non-stationary signals. Such decompositions may provide speech processing and acoustic pattern computation, which can be used by a pattern recognition algorithm. But, thanks to their mathematical foundations, these techniques can powerfully be used as speech feature extraction algorithms. Section 3.5 describes how such techniques are applied to acoustic phonetic decoding, error detection, and control.

A feature vector computed by one of the methods described above is used as the input to the pattern matching stage that is described in the next section. A block diagram of a signal processing scheme based on linear prediction is illustrated in Figure 5.

3.3. PATTERN MATCHING

The pattern matching process consists of comparing the incoming speech with stored representations, which are usually whole-word models but may be phoneme-based. The word model that is most similar to the speech is considered to represent the word spoken. Both the incoming speech and the word models will be represented by sequences of vectors, so to achieve the comparison, one needs some means of measuring the similarity of the vectors and a way of determining which speech vector corresponds to which vector of a model.

The "distance metric" used to measure the similarity between vectors will depend on the signal representation used. The simplest is the Euclidean Distance, i.e. the sum of the squares of the differences between the individual components. Strictly speaking, this is only appropriate if all elements of the vector have the same significance, but factors are usually applied to give most weight to those channels known to carry most information.

In general, the correspondence in time between the vectors of the speech and those of the models is unknown. Even if the times at which a word starts and finishes are known (which is not usually the case), variations in the rate of speaking occur within words. Some speech sounds have relatively constant duration, while others vary widely. It is necessary, therefore, to find the optimum correspondence between the vectors of the incoming speech and those of each model. If the endpoints of the spoken word can be determined, it is possible to use linear time compression, but this is far from the optimum and is only practical for isolated word recognition.

Dynamic Time Warping (DTW) is the simplest means of optimizing the matching between vectors of the incoming speech and those of the models. It is most often used in combination with simple models, such as stored sequences of vectors from single utterances of each word. A detailed description of the algorithm is given in Rabiner and Juang [4]. In outline, a distance score is calculated between each vector of the speech and each vector of the word model. It is then possible to find a sequence of vectors from the model (some of which may be repeated and some may be skipped) which gives the minimum cumulative distance score. This is done using a mathematical technique called Dynamic Programming (or the Viterbi algorithm). The score for each model is normalized to allow for different numbers of vectors. The model that has the lowest score is taken to represent the word spoken.

For years, researchers have been developing Artificial Neural Network (ANN) algorithms, based on models of biological neuron structures (see Figure 6). In speech recognition, the Multi-Layer Perceptron (MLP) is the architecture most commonly implemented. Based on this model, Time Delay Neural Nets (TDNN) were first introduced for speech problems by Waibel et al. [8]. In such a model the basic unit of the neural network is modified, taking into account time delay

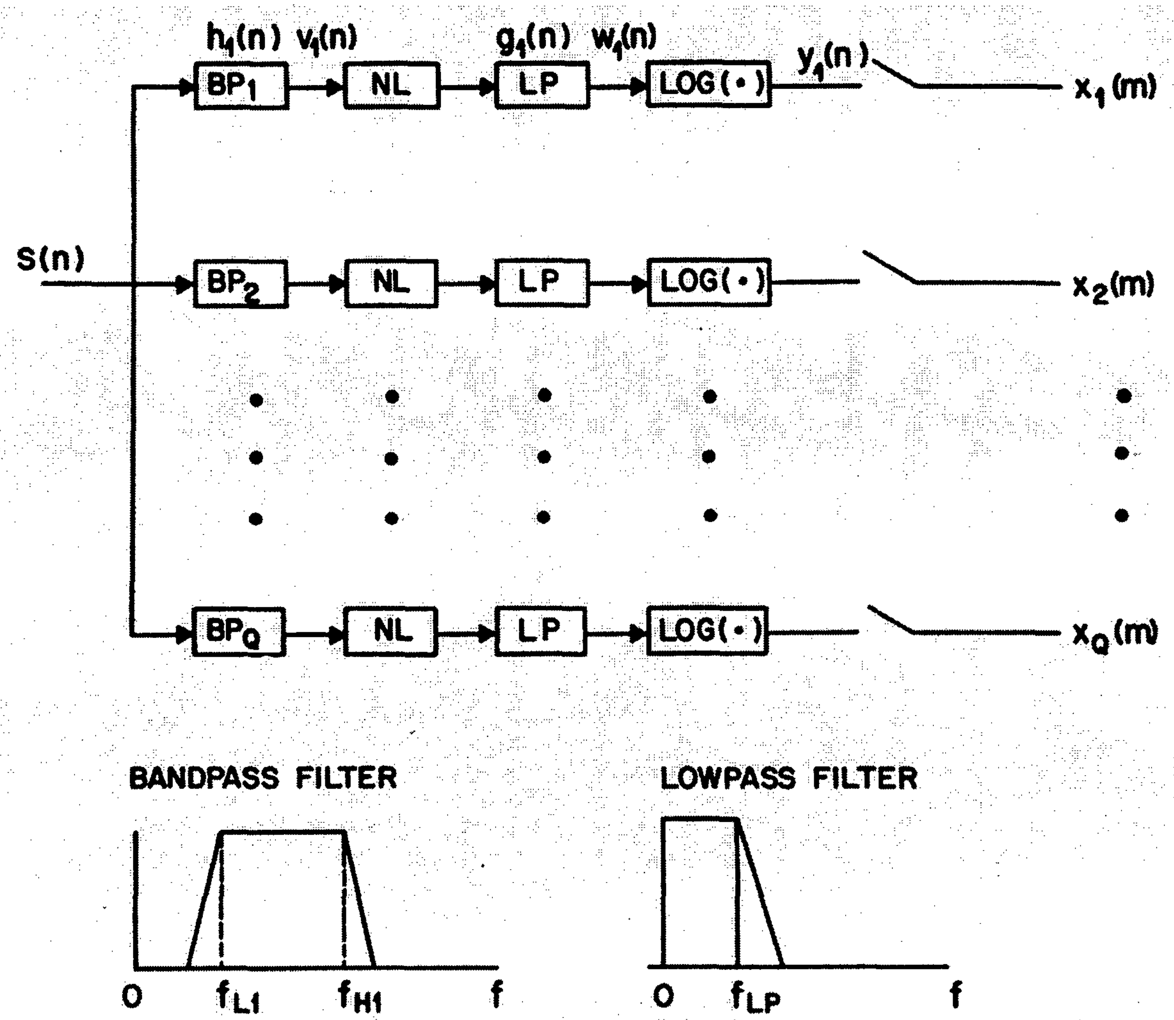


Figure 4 Block diagram of a typical filter bank.

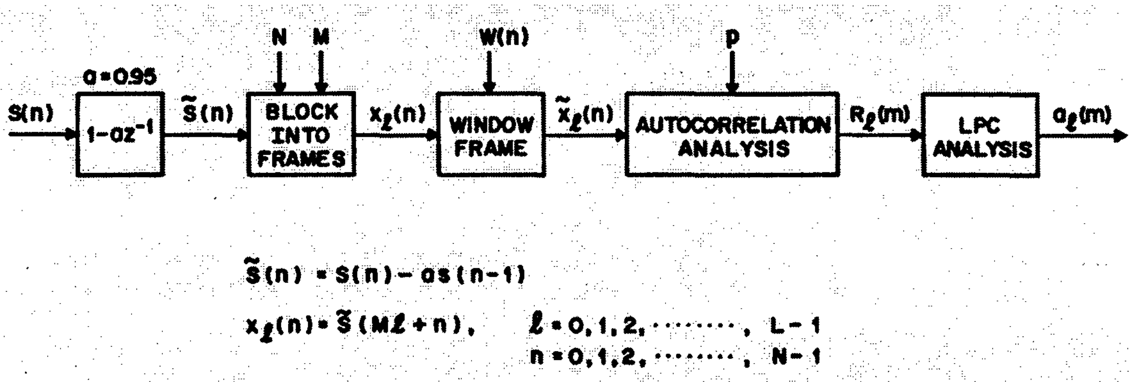
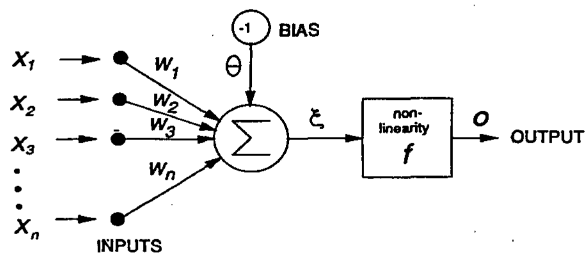
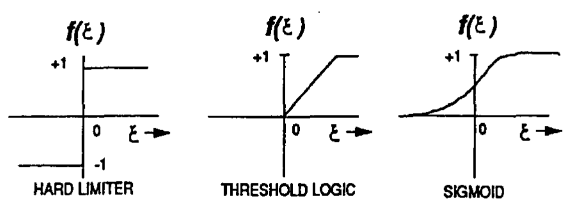


Figure 5 Block diagram of signal processing scheme using Linear Prediction

constraints which are analogous to those used in Dynamic Time Warping.



(a)



(b)

Figure 6 (a) An artificial neuron. (b) Three typical neuron non-linearities, (adapted from [9]).

The most widely used algorithms for pattern matching in ASR today are called Hidden Markov Models (HMMs). In these algorithms, a set of nodes is chosen for a set of phonetic or sub-word units. Five nodes, for example, could represent each phonetic unit [10]. The nodes are connected left-to-right with recursive loops (see Figure 7).

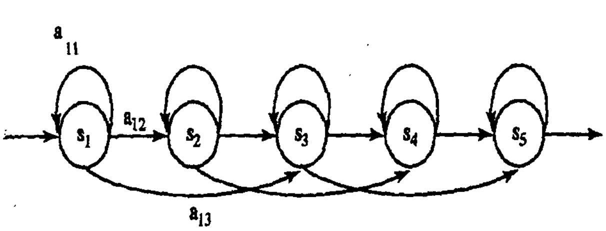


Figure 7 Five-state left-to-right hidden Markov model

Recognition is based on a transition matrix of the probability of changing from one node to another and on a matrix, known as the output probability matrix, representing the probability that a particular set of features (e.g. MFCCs) will be observed at each node. These matrices are generated iteratively during a training process using speech from one or more speakers. These phonetic HMMs are then combined to form larger sets of nodes to represent words. Similarly, the sets of nodes representing words can be combined to form the legal sentences for the particular application.

During pattern matching each HMM model can be used to compute the probability of having generated the sequence of input spectra. This is done very effectively using the Viterbi algorithm [11] on the network of nodes used as the reference patterns. The result of the Viterbi algorithm is the total probability that the spectral sequence was generated by that series of HMMs using a specific node sequence. A different probability value results for every sequence of nodes.

For recognition, the above computation is performed for all possible phoneme models and all possible node sequences. Approximate search algorithms have been used to reduce the search computation without loss in performance. A commonly used technique known as beam search [12] is used to prune nodes that have low probabilities. The one sequence that results in the largest probability is declared to be the recognized sequence of phonemes/words/sentence.

It can be also shown that HMMs and ANNs can be linked together [13]. Such links have led researchers to integrate connectionist networks into a hidden Markov model speech recognition system. Then, it is shown that a connectionist network can be used as a probability estimator: in the classical HMM approach, topologies and probability density functions (pdf) are both chosen, initialized and estimated. In the approach described in [14], the topology of the HMM is still chosen but an MLP is dedicated to the output pdf estimator, through an iterative procedure, alternating between training the MLP and re-estimating the transition probabilities. The efficiency of this method has been shown through an evaluation on speaker-independent databases distributed by the Defence Advanced Research Projects Agency (DARPA). However, this technique remains dedicated to non-noisy speech recognition. Under adverse conditions, embedding preprocessing algorithms should improve their performance (see "Applications of speech-based control", this volume).

3.4. ERROR CORRECTION

It is likely that, for the foreseeable future, speech recognizers will always make some mistakes; after all, humans sometimes mis-hear what is said even under good conditions. In order to provide assurance that the voice input system takes the correct action in response to a spoken command, it is necessary for the user to monitor the recognizer output and have the means to correct any errors that have occurred. Feedback of the recognizer output may take several forms: visual, auditory, or implicit. Where a simple command (two or three words, without digits) is used to perform an obvious action such as changing display modes, no explicit feedback of the recognizer output is required; if the display changes as requested, the command was successfully recognized. If not, it is a simple matter to repeat the command. (There may be a problem regarding what actually did happen as a result of the mis-recognized command.)

More complex or critical commands will require the user to check the recognizer output before the command is executed. Feedback may be visual, (via the head-up display (HUD) or a special display), or auditory. Each has its advantages and disadvantages. Visual feedback is most reliable, but detracts from the eyes-out advantage of voice input. This is somewhat offset by the pilot being able to choose the time at which he looks at the feedback display. Auditory feedback leaves the pilot's eyes free for other tasks, but is transient and may be missed if the pilot's attention is distracted. It may also interfere with, or be overridden by, communications or

auditory warning signals. A study on feedback modality [15] showed that providing both types of feedback gave the best performance on a voice input task and interfered least with a concurrent tracking task.

If an error is discovered in the recognizer output, means must be available to correct it before the command is executed. The simplest way to do this is to delete the whole command and repeat it; this will probably be the most effective way if the error rate is low. Alternatively, the vocabulary, syntax and system interface must provide a means to selectively delete and correct individual words in the command. Words such as "correction," "delete," or "insert" may be used to alter single words or digit strings, in which errors are most likely to occur. However, when the commands consist of only a few words, it is generally easier just to repeat the whole command.

After having decoded possible erroneous speech recognition, a dialogue can be used in order to correct the whole sentence or a single word if the algorithm is accurate enough to localize the possible error inside the sentence. The problem is to design the interface between the man and the machine in such a way that the machine seems simple, or, at least, considerably less complex than it is. Moreover, the dialogue must be as generic as possible in order not to have to design "ad hoc" dialogues from one application to another. So, the problem is to design a generic dialogue core that could be coupled to different applications. Figure 8 summarizes the previous explanations by describing the organization of such a dialogue system.

3.5. ACOUSTIC PHONETIC DECODING

Among all the methods developed during the last decades in speech recognition, one can distinguish "global" methods from "analytic" ones. Global methods recognize utterances by comparison to references, collected through an acoustical model of words. Dynamic Time Warping, Hidden Markov Models and Neural Networks are considered global methods.

Since spontaneous continuous speech production induces coarticulation effects, an analytic approach has been developed in order to localize and identify elementary entities during continuous speech production. According to the set of entities used, one can distinguish Acoustic Phonetic Decoding (APD) where elementary templates are phonemes, diphones, or syllables. Acoustic Features Identification (AFI) localizes and identifies phenomena that occur in speech production through acoustical characteristics such as voiced/unvoiced, plosive or not, fricatives or not, etc. Differences between APD and AFI are small enough to consider them equivalent in this presentation. Even if the analytic approach is a potential method nowadays, global approaches still remain more efficient.

In order to control ASR, we must provide specific algorithms to detect speech recognition errors. One method consists of establishing acoustic phonetic decoding or speech feature extraction (see Figure 8) and analysis to be compared to the solution produced by the ASR. Such an approach is close to the techniques provided in analytic speech recognition, but the goal here is less ambitious than pure recognition: we only want to point out the main features of a sentence through a macro-phonetic classification (voiced/unvoiced speech, voiced/unvoiced fricatives, voiced/unvoiced plosives,...).

Several accuracy levels can be taken into account: for example, if the pronounced utterance is "AUTO" and the ASR solution is "STOP", a voiced/unvoiced classification is sufficient to detect the error. But to separate "four" from "pour", voiced/unvoiced classification is irrelevant and a classification between fricatives and plosives is required. Such an approach could allow the detection of a large portion of speech recognition errors, especially in noisy applications where experiments show that ASR errors are, for the most part, irrelevant from an acoustic phonetic point of view. Such a strategy could not solve some difficult configurations without a perfect classification that would lead to a perfect ASR. But as long as ASR is not perfect, such an approach is relevant. Moreover, for military aircraft applications, such algorithms must be efficient in noisy environments.

As stated in section 3.2, wavelet analysis appears to be a relevant technical method to provide such algorithms. Wavelet decomposition is a powerful tool to analyze short-duration phenomena. After signal decomposition, entropy criteria-based algorithms provide relevant speech segmentation (see [16] and [17]). Moreover, in noisy environments, even in the case of correlated noise, the noise wavelet coefficients tend to be uncorrelated as the resolution and regularity levels increase. Rather than using entropy criteria-based algorithms, another method consists of applying new detection algorithms [18]. These algorithms allow fricatives and plosives detection (see [18] and [19]).

3.6. SPEECH RECOGNITION ASSESSMENT

Speech recognizer performance is often expressed in terms of speech recognition rate. Speech recognition rate must be carefully used. In fact, the connected-word recognizer errors are generally assigned to three categories: deletions, insertions, and substitutions. *Deletion errors* are where nothing in the solution provided by the recognizer matches with a particular word of the utterance. *Insertion errors* are where a word recognized corresponds to nothing in the input. And *substitution errors* are where the word recognized is different from the corresponding word in the input utterance.

Each case is associated with a particular rate and performance is often obtained through a combination of these different rates and can be considered as a Word Recognition Rate (WRR). On the other hand, it is possible to define a Sentence Recognition Rate (SRR) which is computed by considering that the whole sentence recognition is false as soon as there is only one word that has been misrecognized. It is clear that WRR and SRR are quite different. In a military aircraft cockpit, the commonly used Recognition Rate is the SRR. The SRR is more critical because, in aeronautical contexts, speech recognition errors imply consequences on the whole system. So, it appears very important to make speech recognition systems robust in order to avoid critical consequences on the system due a speech recognition error, as mentioned in the previous paragraphs.

4. SUMMARY

In this lecture we have reviewed general framework for speech-based control also know as automatic speech recognition. We have discussed the three stages of speech recognition (signal acquisition, signal processing, and pattern matching) and shown how they contribute to the recognition process. Almost every aspect of continuous speech recognition represents a challenge for aerospace applications of this technology.

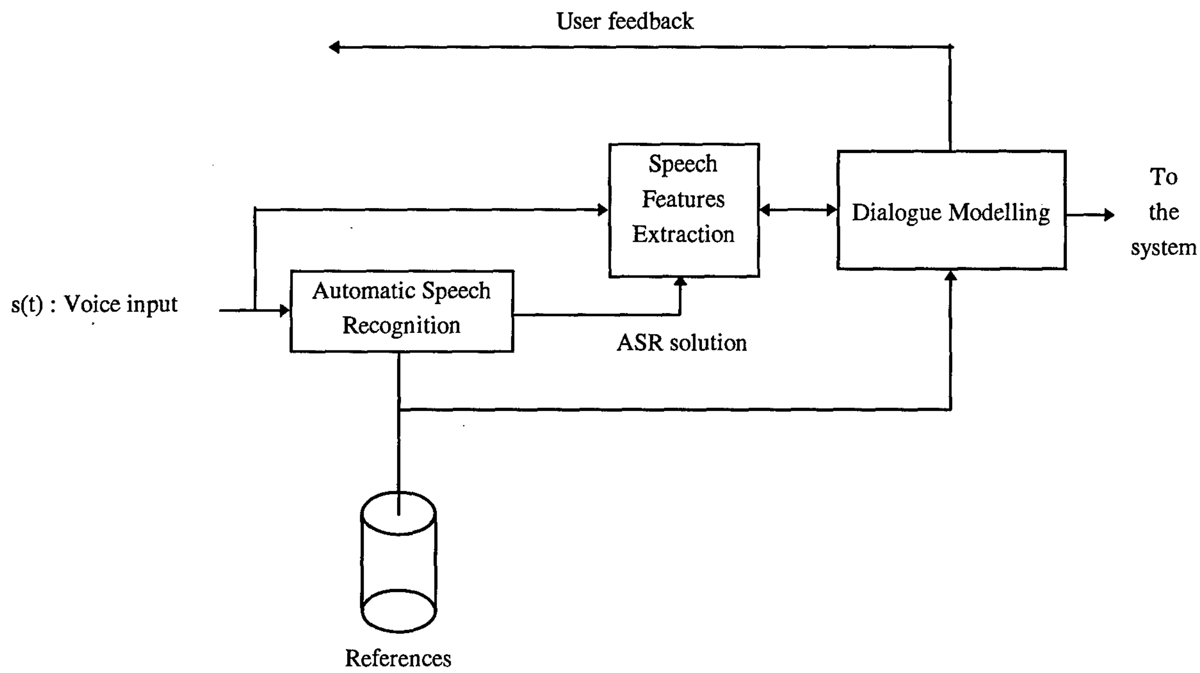


Figure 8 Diagram of Dialogue System

5. REFERENCES

1. Flanagan, F.L., "Speech Analysis, Synthesis and Perception", New York: Springer Verlag, 1972.

2. Deller, J.R., Proakis, J.G., and Hansen, J.H.L., "Discrete-Time Processing of Speech Signals", Englewood Cliffs: Macmillian, 1993.

3. Williams, C.S., "Designing Digital Filters", Englewood Cliffs: Prentice-Hall, 1986.

4. Rabiner, L. and Juang, B.-W., "Fundamentals of Speech Recognition", Englewood Cliffs, NJ: Prentice-Hall, 1993.

5. Hermansky, H., "Perceptual linear prediction (PLP) analysis of speech", J. Acoust. Soc. Am., 87, 4, 1990, pp 1738-1752.

6. Hunt, M.J. and Lefebvre, C., "A comparison of several acoustic representations for speech recognition with degraded and undegraded speech", in "Proc. of Intl. Conf. on Acoust., Speech and Signal Processing", 1989, pp 262-265.

7. Davis, S. and Mermelstein, P., "Comparison of Parametric representations for monosyllabic word recognition in continuously spoken sentences", IEEE Trans. Acoust., Speech and Signal Processing, ASSP-28, 4, 1980, pp 357-366.

8. Waibel, A., Hanazawa, T., Hinton, G., Shikano, K., and Lang, K., "Phoneme recognition using time-delay neural networks", in "Proc. of Intl. Conf. on Acoust., Speech and Signal Processing", 1988, pp 99-102.

9. Lippman, R.P., "An introduction to computing with neural nets", IEEE ASSP Mag., , April, 1987, pp 4-22.

10. Rabiner, L., Levinson, S., and Sondhi, M., "On the application of vector quantization and hidden Markov models to speaker-independent isolated word recognition", Bell Sys. Tech. Journal, 62, 1983, pp 1075-1105.

11. Forney, G.D., "The Virterbi Algorithm", Proc. IEEE, 61, 1973, pp 268-278.

12. Lowerre, B.T., "The Harpy Speech Recognition System", Ph.D. Thesis, Carnegie-Mellon University, 1976.

13. Bourland, H. and Wellekens, C.J., "Links between Markov models and multi-layer perceptrons", IEEE Trans. on Pattern Analysis and Machine Intelligence, 12, 1990, pp 1167-1178.

14. Renals, S., Morgan, N., Bourlard, H., Cohen, M., and Franco, H., "Connectionist Probability Estimators in HMM Speech Recognition", IEEE Trans. on Speech and Audio Processing, 2, 1, Part II, 1994.

15. McGuinness, "Effects of Feedback Modality in an Airborne Communications Task", RAE Technical Report TR 87072, December 1987.

16. Wesfreid, E. and Wickerhauser, M.V., "Adapted Local Trigonometric Transforms and Speech Processing", IEEE Trans. on Signal Processing, 41, 12, 1993.

17. Wickerhauser, M.V., "Adapted Wavelet Analysis from Theory to Software", Boston, MA: A. K. Peters, 1994.

18. Pastor, D., "Diagnostic sur Signaux quasi-atationnaires par Decomposition en Ondelettes Orthonormales et Detection de Coefficients Significatifs", Ph.D. Thesis, l'Universite de Rennes I, 1997.

19. Lemoine, C., "Recherche de traits acoustiques de la parole bruitee par Analyse Multi-Resolution", Ph.D. Thesis, l'Universite de Bordeaux, 1998.

TECHNOLOGY AND APPLICATION OF GAZE BASED CONTROL

Joshua Borah
Applied Science Laboratories
175 Middlesex Turnpike
Bedford, MA 01730

1. SUMMARY

This lecture reviews the potential use of gaze measurement as a means of human interaction with computers and other systems, especially in the military aerospace environment. It addresses the reasons for considering gaze control, reviews techniques for measuring gaze; and discusses physiological, behavioral, and practical considerations for design of gaze based controls.

2. REASONS FOR CONSIDERING GAZE BASED CONTROL

Use of gaze based control is intended to exploit the naturalness, speed and accuracy of visual fixation and tracking. Gaze may be used for explicit control such as designating targets in the external world (E.g., for off-boresight weapons) and selecting items on cockpit displays. It may also be used for implicit control functions such as providing context information for voice or gesture controls, or allowing enhanced resolution of just the local area ("area of interest") being viewed within a display. Gaze control may allow quicker control action while reducing the load on other control resources such as the hands, and in other cases may allow enhanced capabilities (such as the area of interest display) that would not otherwise be possible.

Where line of sight is measured relative to the headgear, it is also necessary to measure the headgear position and orientation in order to compute the eye line of sight with respect to the airframe. In comparison with head pointing alone, gaze based control offers potential advantages of speed, ability to cover a wider angular envelope, and the possibility of less performance deterioration under turbulence-induced vibration or during high-g combat maneuvering.

Many control actions must begin with a visual fixation no matter what the control modality, and there is obvious advantage to exploiting this natural behavior. The natural behavior, however, may not always include a prolonged fixation or an extremely precise fixation, and may include other brief fixations at non related objects. Care is required to maximally exploit natural looking behavior and to minimize requirements for difficult eye gaze actions that must be learned.

3. GAZE TRACKING METHODS

Although mature as laboratory research instrumentation, the current generation of gaze measurement devices has probably not yet reached the level of true practicality for applied use in aerospace cockpit environments. This does, however, appear to be a reachable horizon in the reasonably near term.

Gaze trackers measure line of gaze and/or point of gaze. *Line of gaze* is the imaginary straight line extending from the center of the fovea (the high acuity section of the retina), through the center of the eye lens and out to infinity. *Point of gaze* refers to the point whose image actually forms at the center of the fovea. It is the intersection point of the line of gaze with a visible surface. A gaze measurement system usually includes several subsystems (see Figure 1). An *eye tracker* determines pointing direction of the eyeball with respect to a sensor. If the sensor is head mounted, the system must include a *head tracker* to determine position and orientation of the head with respect to the environment; and if the system is to determine point of gaze rather than just line of gaze, it must also include a processor with knowledge of where visible surfaces are in the environment.

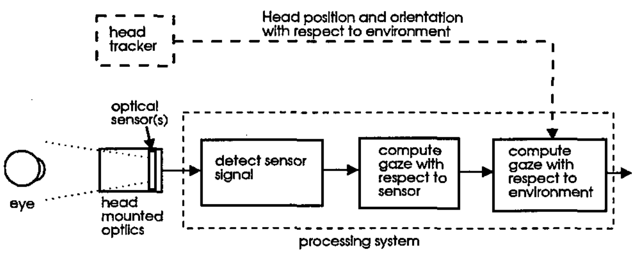


Figure 1. Schematic showing the functional components of a typical gaze tracker

It is important to note that gaze trackers cannot measure what someone is attending to, but rather can only measure aim point of the eye's high acuity area (fovea). This is usually the same as *point of regard* (the part of the visual field that the subject is actually paying attention to), but not always.

Eye tracker performance is often described in terms of the following parameters. *Accuracy* is the expected difference between measured eye line of gaze and true eye line of gaze, usually expressed in terms of visual angle. *Precision* (repeatability) is the expected difference in repeated measurements of the same true eye line of gaze. *Linearity* is the degree to which a change in the measurement is proportional to the actual change in eye angle, and is usually expressed as a percent of the eye angle change being measured. Stated another way, linearity is the amount that a plot of measured values versus actual values is expected to deviate from a straight line. *Resolution* is the smallest change in eye angle that can be reported by the device. *Range* is the amount of eye motion that can be measured, usually specified in degrees visual angle. Range may be specified with respect to the head gear or with respect to the external environment (e.g. airframe), depending upon the device reference frame. *Update rate* is the frequency with which data samples are measured and reported, usually as "samples/second". *Transport delay* is the amount of time that it takes data to travel through the system and become available for use. *Latency* (or *throughput*) usually refers to the amount of time required to accurately reflect a change in the quantity being measured. It is influenced by pure transport delay and also by dynamic operators (for example, a low pass filter) in the

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

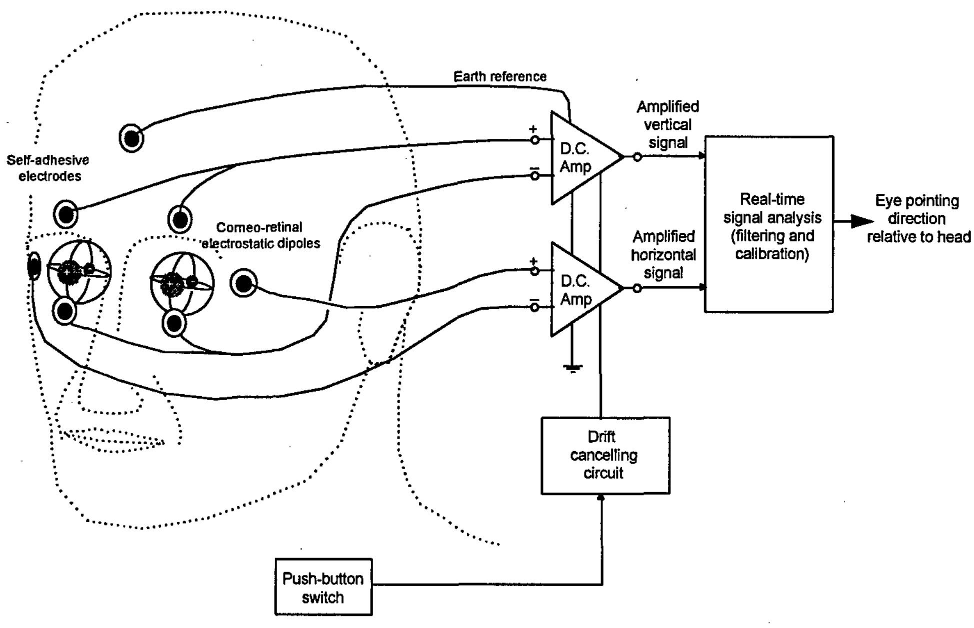


Figure 2. Schematic illustrating electro-oculographic (EOG) technique for measureing eye motion

signal path. *Bandwidth* is the range of sinusoidal input frequencies that can be processed by the system without significant attenuation or distortion.

The component of a gaze measurement system that is currently most critical to achieving practical gaze tracking in operational aerospace environments is the eye tracker, the instrument that attempts to measure the pointing direction of the eye ball. The predominant eye tracking techniques are discussed below. They can be classified as *electro-oculographic*, *scleral coil*, and *optical* methods.

3.1 ELECTRO-OCULOGRAPHY

The retina at the back of the eye develops a small negative electrical charge relative to the front surface of the cornea, probably as a result of its higher metabolism [1]. Electro-oculography uses skin electrode pairs placed on either side of the eye, and above and below the eye to measure the direction of this electric dipole. The corneo-retinal dipole induces zero differential voltage when the dipole axis is about midway between electrodes. A change of about 20 $\mu V/°$ results when the eye is rotated towards one of the electrodes. Rather than making independent measurements for each eye it is common to mount a single pair of electrodes at the outer canthi of both eyes, as shown in Figure 2, to sense their combined horizontal effect. To minimize drift, a "reference" electrode, sometimes sited at the center of the forehead, is usually connected to the amplifier ground.

Small commercial (silver)+(silver chloride) skin electrodes, normally used for monitoring the heart functioning in babies, are commonly employed to minimize electro-chemical artifacts. The skin is cleaned and de-greased with an alcohol swab. Then the contact surface is wetted with conductive saline gel and the electrode is fixed in place using the adhesive backing ring. The short leads are connected to high gain, high (>1 MΩ) impedance, low noise, low drift differential amplifiers having a bandwidth from zero to about 100 Hz. Calibration is required to scale and map the EOG signals to coordinates of gaze with respect to the head. The resulting measurement has high temporal bandwidth and provides an excellent measure of eyeball dynamics, but determination of absolute line of gaze with respect to the head is significantly compromised by drift. In order to keep the signals within the dynamic range of the amplifiers, it is usually necessary to periodically re-zero the output when the subject is known to be looking straight ahead. Extrapolating from laboratory measurements by Shackel [2] and in flight tests conducted in a Jaguar aircraft [3] it seems reasonable to conclude that EOG measures in a cockpit environment might allow inference of eye pointing direction relative to the head with an expected error between about 3° and 7°, assuming some form of filtering and frequent re-zeroing.

3.2 SCLERAL COIL

The scleral coil, first developed by Robinson [4], requires attachment of a sensing element to the subject's eye. A very fine induction coil is embedded in a shallow ring of silicone rubber, the inner surface of which is slightly hollow, so that it adheres to the limbus (the boundary between the clear cornea and the white sclera) by capillary action and suction and remains concentric with the corneal bulge. Thin wires connecting the coil to sensing electronics are usually brought out across the nasal corner of the eye and taped securely to the side of the nose. The subject sits with his head inside two sets of orthagonally oriented Helmholts coils (see Figure 3). One pair of excitation coils creates an oscillating horizontal magnetic field. This field induces a current in the sensor coil proportional to the sine of the horizontal angle between the scleral coil axis and the field. The other set of excitation coils, energized with a signal at 90° phase shift to the first,

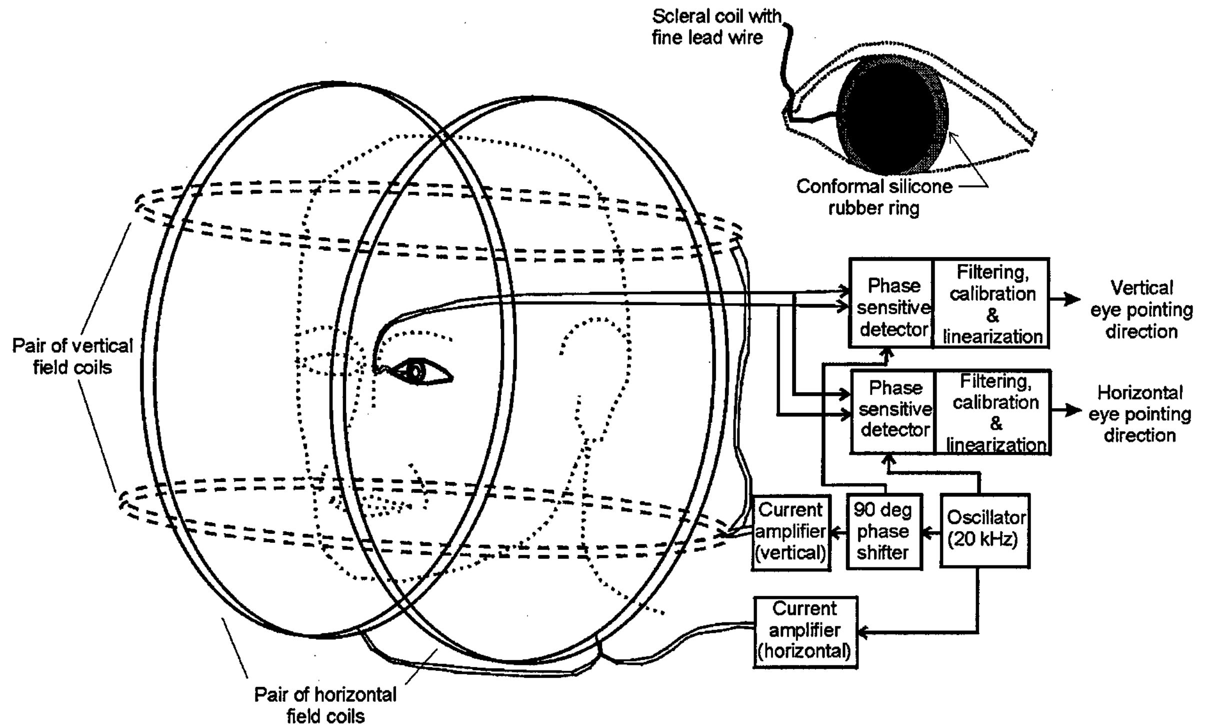


Figure 3. Schematic illustrating the sceral search coil method of measuring eye motion

creates an orthagonal field resulting in induced sensor coil current that is similarly related to vertical angle. Phase sensitive detection is used to find an induced signal that is exactly in phase with each set of excitation coils. If both eyes are outfitted with scleral coils, the pointing direction of both eyes can be measured with respect to the fixed excitation coils. The sclera ring can also be equipped with an orthaginal sensing coil allowing measurement of eye torsion (rotation about the line of sight).

Complete scleral coil systems are available commercially [5]. Scleral coil systems are extremely accurate, fast, and dependable. Following a simple calibration to define the initial reference orientation of the eye, the rotations can be measured to a resolution of about 1 arcmin over a range of about ±15° to an accuracy of about 1% of the range, with a typical bandwidth of 0 to 200 Hz. Scleral coil tracking is also distinctly invasive and requires Helmholz coils that will probably be difficult to integrate on aircrew helmets or affix to the cockpit.

3.3 OPTICAL EYE TRACKING TECHNIQUES

Optical eye tracking techniques make use of optically detectable eye features and geometry to determine the orientation of the eye ball.

The following features, illustrated in Figure 4 are most commonly used :

- *Limbus* -- the boundary between the colored iris and white sclera.
- *Pupil* -- the opening in the iris (aperture of the eye)
- *Corneal reflection (CR), or first Purkinje image* -- mirror reflection of an external source from the outer surface of the cornea
- 4th *Purkinje image (4PI)* – mirror reflection of an external source from the rear surface of the eye lens.

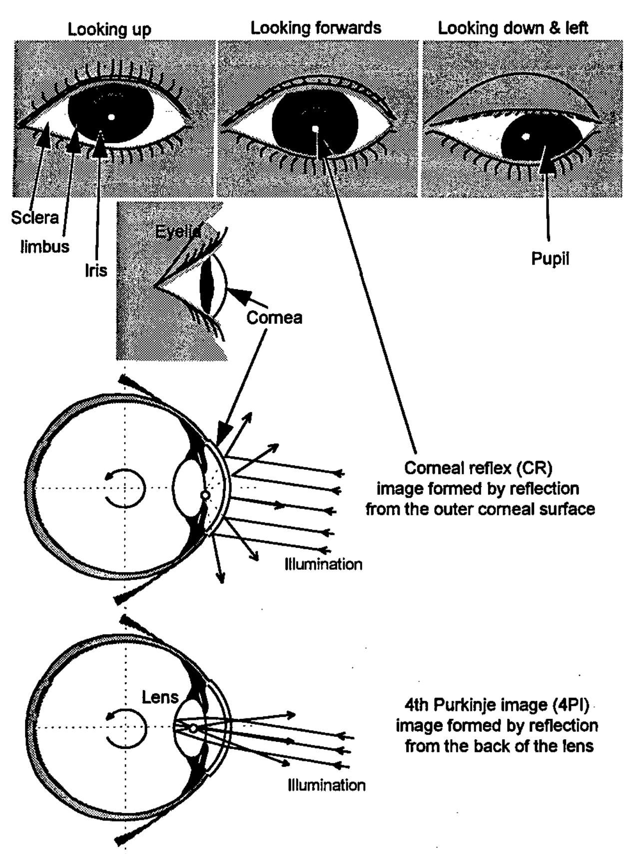


Figure 4. Eye features exploited by optical eye tracking systems

Eye ball orientation can be computed from the position of a single feature if the sensor is assumed to be rigidly fixed to the head or if sensor position with respect to the head can be independently measured.

Position of a single feature alone will not distinguish rotation of the eye ball from movement of the sensor. Multiple landmarks located at different radii from the center of the eye ball will appear to move with respect to one another as the eye rotates, but will move together when the sensor translates. By differentiating between eye rotation and translation with respect to a sensor, dual feature techniques minimize errors due to shifting of head mounted optics, and also allow use of non head mounted optics. Dual feature systems usually use the pupil and corneal reflection, or the corneal reflection (CR) and 4th Purkinje image (4PI). The pupil forms a landmark near the eye ball surface (about 9.8mm from the eye ball center), the CR behaves as would a land mark at the same radius from eye ball center as the corneal center of curvature (about 5.6 mm), and the 4PI appears to move the same amount as the posterior lens surface center of curvature (about 11.5 mm from eye ball center).

Single features or groups of features on the surface of the eye, often dominated by the boundary between the dark iris and white sclera, can be tracked with small numbers of photosensors. An example is shown in Figure 5. This allows very high temporal bandwidth and fine spatial resolution, using essentially analog processing, but results in poor static accuracy because movement of the optics with respect to the eye cannot be distinguished from eye rotation. Very small relative movements are confused with relatively large rotations.

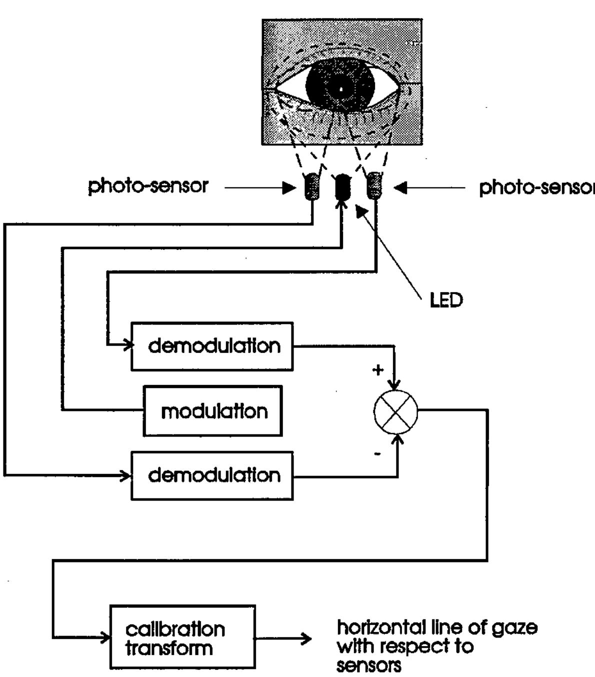


Figure 5. Schematic showing a simple photo sensor system for tracking horizontal position of the limbus. As the eye rotates, proportionately more white sclera is exposed to one photo-sensor than to the other.

The relative positions of the corneal reflection and 4th Purkinje image ("dual Purkinje image technique") can be used to track eye rotation with excellent precision and accuracy (on the order of 20 arc seconds). The only commercially available dual Purkinjie image system images the two features onto separate quadrant detectors, and uses separated closed-loop servo-controlled pathways to keep the features centered on the detectors. The servo-control signals are a measure of the feature positions. This is also an essentially analog process allowing very high temporal bandwidth.. Range, however, is relatively small (typically ±10°) because the 4PI is visible only within the lens opening and is obscured when it moves behind the iris. The only currently available system is engineered as a large bench mounted optical assembly that is impractical for airborne application.

The most prevalent optical technique, and the one used by systems that currently come the closest to being practical for airborne use, tracks the relative position of the corneal reflection and the center of the pupil. CR/Pupil devices are not as precise or accurate as the dual Purkinjie image device, but can be far less obtrusive and easier to operate primarily because the pupil is easier to detect the 4PI.

Generally the eye area is illuminated by a near infra red source (or multiple sources) and a solid state video camera captures an image of the eye. The camera is typically filtered to receive only light of the wavelength produced by the eye tracker's near infra red source.

If the optics (camera, illuminator, and lenses) are mounted to the user's head gear, a hot mirror (beam splitter that reflects IR and transmits visible wavelengths) is usually used to reflect near IR light to the optics while still allowing unobstructed vision for the wearer. This is illustrated is Figure 6. Alternately, non head mounted optics may use a moving mirror or moving camera platform to follow head motions.

The eye acts as a retro-reflector. If the eye illumination beam is coaxial with the camera, light reflected back from the retina is captured by the camera making the pupil appear to be a bright circle. This accounts for the red eye effect sometimes produced by flash photography. Off axis illumination produces the more familiar dark (black) pupil image (see Figure 6). Dark pupil images provide easier pupil detection in very bright environments (e.g. sunlight), whereas bright pupil images provide easier detection in darker environments.

The corneal reflection is significantly brighter than any other visible feature, including a bright pupil image, and is relatively easy to detect so long as it is not obscured by the eye lids or confused with corneal reflections from some external sources.

Real time image analysis is used to identify the pupil and corneal reflection and find their centers. Relative feature brightness is often a primary discrimination criterion, but more and more sophisticated pattern recognition techniques are being used as the amount of readily available digital processing power increases. This makes it possible to recognize the features of interest in real conditions and cope with extraneous reflections, partial eye lid occlusion and motion-induced blur [6]. Calibration is required to account for individual eye geometry and optics placement.

Range for CR/Pupil systems is limited to about ±25° by CR excursion within the cornea for systems with a single illumination source, but can be extended considerably if multiple sources are used, especially on the horizontal axis. Vertical range in the downward direction is often limited, by

eye-lid occlusion of the pupil, to 5-10° below the line of gaze that would look directly at the eye camera. For this reason the optics are usually set so that the camera views the eye from about 5° below the nominally straight ahead direction. Note that for head mounted optics these range limits apply to eye motion with respect to the head. Measurable line of gaze with respect to the airframe is limited only by head motion.

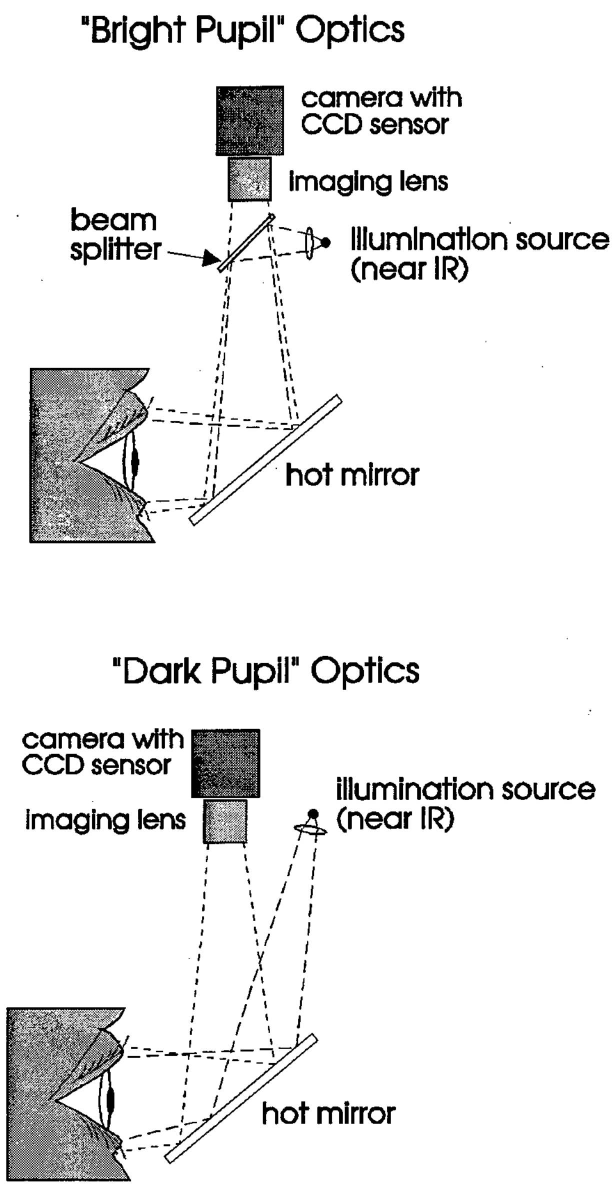


Figure 6. "Bright pupil" and "dark pupil" optics for CR/pupil tracking

Accuracy is usually on the order of 1° visual angle for CR/Pupil systems. It may vary from 0.5° or better during very careful laboratory tests with selected subjects, to > 2° under difficult conditions or with difficult subjects. Precision and resolution are usually in the range of 0.1° to 0.5° depending upon the particular system and upon operating conditions. Frequency response is usually limited by video frame rates to 50 or 60 samples/sec. Higher update rate sensors are available, but require sacrifices in spatial resolution, physical size, and sensitivity.

There are numerous developed CR/pupil systems, but currently no militarized systems, and no systems robust enough for operational military use. Current systems have difficulty handling bright sunlight which tends to saturate the entire camera image and degrades effective contrast, are not able to automatically make adjustments for changing ambient conditions, have not yet been properly integrated with military aviation head gear, and are not quite reliable enough for operational service. It is reasonable, however, to think that the necessary advances to CR/pupil systems can be made in the near term.

Illumination of the eye by optical eye trackers must, of course, remain within safe levels. Safety standards are published by numerous agencies including *The American Conference of Governmental Industrial Hygienist (ACGIH), The American National Standards Institute (ANSI)*, the *Federal Food and Drug Administration* (FDA), and *The International Electrotechnical Commission (IEC)*. The standards vary, with the IEC standard currently being the most restrictive. Under IEC standards, for example, the source must be safe even if viewed from the closest mechanically possible distance through a magnifying glass.

3.4 CALIBRATION

All eye tracking methods require a transformation to convert the measured quantity to the desired quantity. For example separation between the pupil and CR must be converted to point of gaze coordinates on a surface, or a line of gaze vector in a particular coordinate frame. For all eye tracking methods discussed, with the possible exception of the scleral coil, the transformation parameters vary between subjects, optics placement, and other conditions.

Calibration refers to a procedure for gathering data, and for using the data to compute transformation parameters. The procedure usually consists of asking the subject to look at number of pre-defined points, while storing samples of the measured quantity (e.g. pupil and CR position).

The transformation can either be a form of interpolation, a set of continuous equations, or some combination of these. The details vary widely among available systems. Theoretically, the transformation can remove any systematic error that is a function of the measured quantities. In practice, there is a limit to the amount of data that it is reasonable to gather with a calibration procedure, and therefore a limit to the complexity of the transformation.

The accuracy and linearity of eye tracker measurements are limited by the underlying precision (repeatability) of the measured quantity. Up to that limit, accuracy and linearity are determined by the quality of the calibration and transformation scheme. Adding calibration target points, and using the additional data to add interpolation points or to increase the order of a polynomial transform, usually improves accuracy, but with diminishing returns. Typical calibration schemes require 5 or 9 pre-defined points, and rarely use more than 20 points.

An example of a 2 dimensional interpolation scheme can be found in references [7] and [8]. A cascaded polynomial curve fit method is described in reference [9]. Many other variations of these schemes are possible.

3.5 PROGNOSIS FOR AIRBORNE OPERATIONAL USE

Eye tracking is a relatively mature technology only in the R&D domain. No currently available eye tracking systems are dependable enough or automatic enough for operational

flight applications, nor are there any current systems available in a militarized configuration. All current devices require a skilled equipment operator (other than the person being measured) for optimal use.

Scleral coil and differential PI tracking seem likely to remain laboratory techniques unless some major leaps in the technology occur. These techniques are by far the most accurate of the major techniques in use, but they both present major practical problems. Scleral coil tracking is invasive and requires a Helmholz coil that will probably be difficult to integrate on aircrew helmets. Differential PI tracking has too restrictive a range limitation and too complex an optics package to be easily helmet mounted and ruggedized.

EOG may very well have a place in aircraft as a back up measurement system, an enhancement to add temporal bandwidth to another type of eye tracker, or for use in some control function that requires only detection of eye movement, rather than absolute line of gaze. The accuracy of EOG alone is never likely to be adequate for line of gaze determination.

Differential CR/pupil tracking systems are generally the most unobtrusive eye tracking systems available, and, with head mounted optics, currently come closest to being appropriate for operational use in flight. Those systems, using dark pupil optics along with some form of illuminator strobing and sensor shuttering, are currently best able to operate in daylight and under vibration. The static accuracy (about 1° visual angle) and range (50° horizontal and 40° vertical field with single illumination source) of current CR/pupil tracking devices is adequate for implementing or assisting a variety of tasks in the aerospace environment, although increased accuracy would certainly expand the potential use for eye tracking.

No currently available CR/pupil systems are yet nearly robust enough, automatic enough, or properly integrated with aircrew head gear and military electronics. Robustness must be significantly improved to ensure dependable operation for different users under varying light conditions in operational environments. Automatic operation must be significantly enhanced so that the user can don the equipment and calibrate the system without help, and there-after depend upon proper operation with no intervention by a second person. Optics must be integrated with the appropriate head gear and head mounted display systems, and both optics and electronics must be hardened and militarized. Work is underway in all of these areas, and there is no reason to think that such enhancements cannot be realized with currently available optics, sensor, and processing technology.

Table 1. Summary of most prevalent eye tracking techniques

Method	Typical Applications	Typical Attributes	Typical Reference Frame(s)	Typical Performance
• EOG	• Dynamics of saccades • smooth pursuit • nystagmus	• High bandwidth • Eyes can be closed • In expensive • Drift problem (poor position accuracy) • Requires skin electrodes - otherwise unobtrusive	• head	• static accuracy: ~3°-7° • resolution: with low pass filtering & periodic re-zero, virtually infinite • bandwidth: ~100 Hz
• Scleral Coil	• Dynamics of saccades, smooth pursuit, nystagmus • Miniature eye movements • Point of gaze • Scan path	• Very high accuracy and precision • Invasive • Very obtrusive	• Room	• accuracy:~15 sec arc • resolution: ~1arc min. • range: ~ 30° • bandwidth: ~200 Hz
• Limbus (using small number of photo sensors)	• Dynamics of saccades, smooth pursuit, nystagmus • Point of gaze • Scan path	• High bandwidth • Inexpensive • Poor vertical accuracy • Obtrusive (sensors close to eye) • Head gear slip errors	• head gear	• accuracy: varies • resolution: 0.1° (much better res. possible) • range: ~30° • update rate: 1000 samples/sec
• CR/Pupil	• Point of gaze • Scan path	• Minimal head gear slip error • Unobtrusive • Low bandwidth • Problems with sunlight	• Head gear • Room (airframe)	• accuracy: ~1° • resolution: ~ 0.2° • hor. range:~50° • vert. range: ~40° • update rate: 50 or 60 samples/sec
• CR/4PI	• Dynamics of saccades, smooth pursuit, nystagmus • Miniature eye movements • Point of gaze • Scan path • Image stabilization on retina • Accommodation	• Very high accuracy and precision • High bandwidth • Obtrusive (large optics package, restricted head motion) • Limited range	• Room	• prec: min. of arc • range: ~20° • update rate: 500 samples/sec

Significant improvement of CR/pupil system accuracy is clearly possible, but far less certain, especially in operational environments. Improvements can theoretically be made by using increased processing power to more accurately find the center of the oval pupil in the presence of partial image occlusions, and ragged edges; use of higher order calibration schemes to remove more of the systematic error; use of additional variables in calibration, such as pupil diameter, to further account for systematic effects; use of precision sensor arrays, or sensor arrays that are mapped and compensated for spatial non-linearity's; etc. Such gains may be more than counter-balanced, however, by additional error introduced under the rigors of operational use. Furthermore, the lengthy, careful calibration procedures probably required for exquisite accuracy may be contrary to operational imperatives for quick and easy set-up. Designers may want to consider eye tracking tasks tailored to require less rather than more accuracy and precision in order to improve the chances of meeting robustness and ease of use requirements.

Major eye tracking techniques are summarized in table I

4. HUMAN PHYSIOLOGICAL AND BEHAVIORAL CONSIDERATIONS

Even if instrumentation could make perfect point of gaze and line of gaze measurements, human physiological and behavioral characteristics impose certain constraints, and it is very important to keep these characteristics in mind when considering gaze based control.

Scanning behavior is described by a series of fixations (stopping points), saccades (extremely rapid jumps between fixation points) and smooth pursuits. When examining a stationary scene, both lines of sight are simultaneously held steady for short periods (usually 200 - 600 msec), called fixations, to bring a feature of interest within the approximately 1° angular range of the fovea. Miniature eye movements, of up to several minutes of arc do occur during the periods of "fixation", but are not perceived.

Rapid jumps, called saccades, move the eye between fixations.. Saccades usually reach velocities of 400-600 °/sec, and last 30 - 120 msec. Vision is significantly suppressed during this period. Although saccades can be as large as 50°, they are more commonly 1 - 20°. If a target appears in peripheral vision, it takes a minimum of about 100 msec for a saccade towards the target to be initiated.

When observing a slowly moving object the lines of sight usually track smoothly, but this pursuit reverts to fixations and saccades when the object moves faster than about 30°/sec. Without specific training, smooth eye movements are only possible when following a smoothly moving target or compensating for head movement. A thorough review of eye movement behavior can be found in [10].

Even if we could measure direction of the visual axis perfectly we would not have perfect knowledge of point of regard. Visual acuity is best on the foveal region of the retina, and within the fovea is best near the very center. People therefore direct the visual axes of the eyes (axes passing through the center of each eye lens and fovea) to objects that they want to see clearly; however, there may be a foveal "dead zone" or "indifference threshold" on the order of about 0.3 degrees visual angle for fixation of stationary targets [11]. Attention can be shifted within the foveal region, and even outside of the foveal region if the target of interest falls within acuity limits of peripheral vision [12, 13, 14]. Furthermore, foveation accuracy falls off markedly if a person attempts to maintain fixation for several seconds [15], when tracking a moving target [10, 16, 17, 18], or during rapid head movements [19]. If a person consciously attempts to fixate a small, stationary, target, for a short time, while holding their head steady, we can probably assume the visual axis to be within 0.3 degrees of the target. This should not be confused with visual acuity. People can visually determine whether one object is aligned with another (a traditional aiming task) with precision on the order of arc minutes.

Eye movement with respect to the head (rotation of the eye ball within the socket) has a maximum range of about ±50° horizontally, and about +40°, -60° vertically , but normally remains within about ±15°-20° [20, 21]. Gaze shifts beyond the central 20° field are usually, although not always, accompanied by head rotation. Horizontal eye rotation with respect to the head of more than about 40° from the central position becomes quite uncomfortable if maintained for several seconds.

The normal fixation/saccade pattern of visual scanning can be thought of as a continual series of snap shots that are used to create a mental image of the visual environment; however this is usually an unconscious process. Perception of the environment is of the "single picture" formed in the brain. People are normally not aware of their scan pattern (although they can be if they make a special effort), but rather are normally aware only of the resulting mental image of the environment.

People are not accustomed to consciously controlling their gaze. Unintentional actions initiated by unintentional glances is sometimes called the "Midas touch" problem of gaze control.

It is difficult and annoying, although possible, to maintain steady fixation on a single target for significantly more than a second. Fixations of several hundred milliseconds are most natural. There is also a strong tendency to make quick glances at other nearby targets during unnaturally long fixations.

The eye is drawn to features, and it is very difficult to fixate a blank spot.

Secondary visual feedback (presentation to a person of their own gaze point as measured by an eye tracker) must be handled carefully. Continuous feedback of measured gaze position, if not very accurate and up to date, can sometimes be distracting instead of helpful. If the displayed indicator is slightly displaced from the central line of gaze there may be a tendency to continually try to look at it, leading to a positive feedback loop; however, it requires only minimal conscious effort to avoid this.

The vestibular system helps the brain to stabilize the perceived visual field in the presence of head motion and vibration, and to partially, but not completely, stabilize line of gaze with respect to the visual environment. Steinman and Collewijn [19] showed, for example, that in the case of voluntary head rotations of 2.5-5 Hz during fixation on a distant target, eye motion sometimes compensates for only about 80% of head motion and the two eyes do not move equally, although vision remains clear and fused.

Eye movement appears to remain unaffected under G_Z loading that is sufficient to make head motion difficult. This is supported by only a small amount of empirical data [22], but is consistent with anecdotal evidence and mechanical analysis. The eye is well supported, and has a relatively small moment of inertia. Because of its roughly spherical, homogeneous structure inertial forces would not be expected to produce large rotational moments. Furthermore, eye movements do not cause the disorienting motion sensations that can be produced by moving the head, and hence the vestibular sensors, in the presence of high inertial forces.

5. DESIGN CONSIDERATIONS FOR GAZE BASED CONTROL

5.1 FIXATION FILTERING

As previously discussed Scanning behavior is described by a series of fixations, saccades, and smooth pursuits. When using gaze measurement to determine point of regard, as often required for gaze based control functions, it may be desirable to recognize fixations while filtering out saccades, pursuits, and measurement noise. This is usually done by looking for periods of longer than some threshold time during which gaze remains within a threshold (maximum) area or during which eye movement velocity remains below a threshold. Typical threshold values would be 100 msec, 1° visual angle (circular radius), and 10°/sec, respectively.

5.2 USE OF EYE POSITION FEEDBACK

It is usually important to provide some sort of feedback so that the user of a gaze control system knows that the system is functioning properly and can make adjustments if necessary. In its simplest form this feedback is a cursor superimposed on the visual scene continually showing measured gaze point In theory such feedback, often called secondary visual feedback, can improve system accuracy by allowing the user to adjust his gaze point to correct measurement errors. There is empirical evidence that precise secondary visual feedback can improve visual smooth pursuit performance [23, 24]. In practice, this type of feedback may sometimes prove more annoying than useful depending upon the amount of error, noisiness, and latency in the gaze measurement [25, 26, 27]. In such cases it may help to present measurements that have been filtered by a fixation detection algorithm (feed back cursor moves only when a new fixation is detected), and/or to display a transparent cursor that is at least as large as the expected gaze measurement system error. Alternately it may sometimes be best to present feed back by highlighting the object that the system computes gaze to be indicating, for example, a display icon or object, rather than by continually displaying measured gaze point. If the gaze control task is to designate an external target such as an opposing aircraft or ground target, then some form of continual measured gaze point feed back may be required.

5.3 ACTION CONFIRMATION

Even if gaze is available as a control input, not all fixations will be intended as explicit control actions. There will still be present the semi conscious pattern of fixations ("snap shots") that normally form our perception of the visible environment. To avoid a "Midas touch" affect, any explicit gaze control system must include means to differentiate fixations that are intentional control actions from glances intended only to acquire visual information. Confirmation protocols can range from requiring a slightly unusual gaze behavior (longer than average fixations, a sequence of blinks, or a long blink) to manual action. Control actions that are more consequential or more difficult to undo should require a more reliable mode of confirmation than less consequential actions. Citing an example from Jacob [25], if gaze information is used to cause menus on a display to automatically "pull down" when the menu title is fixated, and "roll up:" again when the menu is no longer being viewed, the consequence of unintentional activation of this feature is minimal. A sensible activation criterion might be a fixation on the menu title that is slightly longer than the average 250 msec fixations that characterize scanning behavior. Further increasing the required fixation time makes unintentional activation less likely at expense of longer task execution times and increasingly unnatural behavior.

An example at the other extreme would be use of gaze to designate external targets to a weapon delivery system. In this case very high precautions must be taken against unintentional activation, and manual confirmation, in the form of a trigger pull or button press is probably warranted. Furthermore, the pilot must first have good feedback providing assurance that the intended target has indeed been selected.

Jacob [25] used a hierarchical set of techniques to confirm display manipulation actions depending on the consequences of inadvertent action. Actions that were benign and easily reversible required only short fixations for activation. Actions that were not as easily reversible required longer fixations or manual confirmations. He found that when the eye tracker was working accurately and dependably it felt as though the system were "reading the user's mind", but when eye tracker performance was not stable enough or not accurate enough it was extremely frustrating to the user.

5.4 GAZE CONTROL TASKS

Explicit gaze control tasks usually involve some type of gaze designation, including designation of external targets from within a cockpit, selection of real or virtual switches within a cockpit, and a range of "cursor control" type tasks (designation of icons, menu labels, screen locations, etc.) associated with fixed or head mounted displays.

Unambiguous gaze designation requires that targets be separated by at least twice the maximum expected error (E.g., 95% confidence interval) of the measurement, so that confidence intervals drawn about adjacent targets will not overlap. With current state of the art for unobtrusive eye trackers, this would correspond to separations of at least 2° visual angle (to accommodate 1° errors), and probably somewhat more in demanding environments. If gaze measurement is to be used to designate an external target to a "lock before launch" weapon system, the capture field of the "locking" system must correspond to a similar visual angle. If icons on a display are to be designated by gaze alone, the same limit (twice the expected error) defines the minimum space between the borders of adjacent icons, or at least between central fixation targets within each icon.

Eye designation has been investigated in the lab, and generally has been found to be as fast or faster than manual designation, and faster than head designation, so long as the eye tracker is working dependably and so long as the task

employs large enough targets to be well within the accuracy capability of the eye tracker [26, 27, 28, 29, 30]. Applied use of eye designation to date has been primarily restricted to systems that facilitate communication for people with motor control disabilities, and there are several commercially available systems that specifically support this application. Performance of eye designation tasks may sometimes be significantly enhanced by use of fixation filtering algorithms, but in general, accuracy of unobtrusive eye tracking systems does not permit as fine a control capability as mouse, trackball, or other manual techniques.

Multi mode control can be used to achieve greater precision, again at the expense of longer task execution times and a requirement for less natural behavior. For example, gaze can be used to position a display cursor near the point of gaze, followed by use of a manual control for fine positioning. Especially in the case of a large cluttered display it is sometimes time consuming simply to find the cursor. In this case, use of gaze to quickly position the cursor to within easy view has the potential to save time over manual control alone; although it would not be as fast or natural as designating a large enough target with gaze alone. Note that besides the manual (or other mode) control for fine cursor positioning, the multi mode technique requires a switch to designate the mode change, and some learned behavior to properly sequence actions (*gaze > mode switch > manual control*).

Gaze measurement can be an excellent tool for "context disambiguation" of voice commands. For example in response to the verbal command *"zoom"*, point of gaze can be used to determine which display or display section to expand. Citing an example from a mission planning interface developed by Hatfield [29], the *command "nav designate steer point"* sets the steer point to the radar display position being fixated at the time the verbal referent *"designate"* was detected. In this way some operations that would otherwise be sequential can be made concurrent. Although context disambiguation takes advantage of natural behavior (E.g., a person is likely to be looking at the display that they want to "zoom") this type of control is not entirely implicit. A particular behavior is required which, although likely, is not certain without explicit intent. Use of gaze to provide context and position information for verbal commands has been tested fairly successfully in several laboratory studies [25, 31, 32, 33, 34].

Gaze measurement has been used in military aircraft simulation to create computer generated, out-the-window displays which have high resolution only in the immediate area about the pilot's gaze point. This "area of interest" display technique was originally motivated by the difficulty in producing computer graphics with both the wide field and high resolution desired. Since high visual acuity is present only in the area near the gaze point, only this area need have high resolution at any given time in order for the entire display to be perceived as being rich in detail. In future operational cockpits, a similar moving area of interest concept might prove valuable with respect to head mounted displays in order to increase the richness of the display about the region of the current gaze point. This may, for example, involve slaving an external sensor with a very narrow cone of sensitivity to follow gaze, and displaying resulting information in the corresponding field within a head mounted display.

Use of gaze measurement for external sensor slaving and/or area of interest display function constitutes an implicit control function. It works entirely in response to normal gaze behavior. For the simulator application described above it proved important to minimize latency and to smoothly blend the boundary between high and low resolution sections in order to preserve the illusion of a detailed out-the-window world. Total latency (gaze tracker plus display) probably needs to remain below 50 msec to preserve the illusion, and this proved a significant problem for the simulation application. Future operational application of this concept is more likely to have the purpose of making additional information available rather than creating an illusion of realism, and although the blending and latency factors will still be of importance in order to avoid annoying or distracting the user, the criteria may be somewhat less stringent.

6. CONCLUSIONS

Gaze measurement may enable a range of potentially useful explicit and implicit control functions. The technology is not yet mature enough for operational airborne application, but the necessary advances can probably be made in the near term. When considering gaze control functions, designers need to consider natural human gaze behavior and the tradeoffs between performance requirements imposed (precision, accuracy, latency, etc.) and robustness, simplicity, and ease of use in difficult environments.

7. AKNOWLEDGEMENTS

This presentation is based on work done as part of AGARD Working Group 25, and contains contributions from all members of that group. It relies especially heavily on the efforts of two other members of that group: Dr. Don Jarrett, DERA, Farnborough, U.K,; and Dr. Alain Leger, Sextant Avionic, Bordeaux, France.

8. REFERENCES

1. Young L. R. and Sheena D., "Eye Movement Measurement Techniques", in Webster (Ed) "Encyclopedia of Medical Devices and Instrumentation", New York, John Wiley & Sons, 1988

2. Shackel B., "Eye movement recording by electro-oculography", in "A manual of Psychophysiological methods", North-Holland Publ. Co., 1967.

3. Vivash J. P., Belyavin A. J., "Eye movements under operational conditions", in Waters M. and Stott J. R. R. (Eds) "Journal of Defence Science", 1, 2, 1996.

4. Robinson D. A., "A method for measuring eye movement using a scleral search coil in a magnetic field", IEEE Transactions on Biomedical Electronics, BME-10, 1963, pp 137-145.

5. Ferman L., Collewijn H., Jansen T. C. and Van den Berg A. V., "Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral coil technique", Vision research, 27, 1987. pp 818-828

6. Brinicombe, A. M., Boyce, J. F. and Durnell, L., "Direction of regard determination", in Delogne, P., (Ed)

"Proc. Intl. Conf. on Image Processing", Lausanne, Switzerland. IEEE Signal Processing Society, 1996.

7. McConkie, G. W., "Evaluating and reporting data quality in eye movement research", Behaviu Research Methods & Instrumentation, 13, 2, 1981, pp 97-106.

8. Kliegle, R. and Olson, R. K., "Reduction and calibration of eye monitor data", Behavior Research Methods & Instrumentation, 13, 2, 1981, pp 107-111.

9. Sheena, D. and Borah, J., "Compensation for some second order effects to improve eye position measurements", in Fisher, D. F., Monty, R. A., and Senders J. W., (Eds) "Eye Movements: Cognition and Visual Perception", Hillsdale, Lawrence Erlbaum Associates, 1981.

10. Hallet,P.E., "Eye Movements", in Boff, K. R., Kaufman, L., and Thomas, M. P. (Eds) "Handbook of Perception and Human Performance - Vol. 1", New York, John Wiley and Sons. 1986.

11. Young, L.R., "The sampled data model and foveal dead zone for saccades", in Zuber B. L. (Ed) "Models of Oculomotor Behavior and Control", Boca Raton, CRC Press, 1981.

12. Julez, B., Gilbert, E. N., Shepp, L. A., and Risch, H. L., "Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited", Perception, 2, 1973, pp 391-404.

13. Bergin, J. R. and Julez, B., "Parallel versus serial processing in rapid pattern discrimination" Nature, 303, pp 696-698.

14. Scinto, L. F. M., "Retinal inhomogeneity and the allocation of focal attention during fixation", in "The Annual Meeting of the Applied Vision Association" St. John's College, Oxford. 1988.

15. Yarbus A. L., "Eye Movements and Vision", New York, Plenum Press, 1967.

16. St.Cyr G. L. and Fender, D. H., "Non-linearities of the human oculomotor system: Gain", Vision Research, 9, 1969, pp 1235-1246.

17. Michael, J. A., Melvill Jones, G., "Dependence of visual tracking capability upon stimulus predictability", Vision Research, 6, 1966, pp 707-716.

18. Kowler, E. and McKee, S., "Sensitivity of smooth eye movement to small differences in target velocity", Vision Research, 27, 6, 1987, pp 993-1015.

19. Steinman, R. M. and Collewijn, H., "Binocular Retinal Image Motion During Active Head Rotation", Vision Research, 20, 1980, pp 415-429.

20. Yamada. M., "Head and eye coordination analysis and a new gaze analyzer developed for this purpose", in d'Ydewalle G. and Van Rensbergen, J. (Eds) "Visual and oculomotor functions", Elsevier, 1994, pp 423-434.

21. Bahill, A. T., Adler, D. and Stark, L. "Most naturally occurring human saccades have magnitudes of 15 degrees or less", Investigative Ophthalmology, 14, 1975, pp 468-469.

22. Sandor, P. B., Hortolland, I., Poux, F., and Leger, A., "Orientation du regard sous facteur de Charge Aspects methodologiques Resultats preliminaires", in "AGARD Meeting on Virtual Interfaces: Research and Applications", October, 1993.

23. Peli, E. and Zeevi, Y. Y., "Multiple visual feedback loops in eye movement control". in "XII International Conference on Medical and Biological Engineering", Jerusalem, 1979.

24. Kenyon, R. V., Zeevi, Y.Y., Wetzel, P. A., and Young, L. R., "Eye movement in response to single and multiple targets", US Air Force report AFHRL-TR-84-29, 1985.

25. Jacob, R. K., "Eye Tracking in Advanced Interface Design", in Barfield, W. and Furness, T. A. (Eds.) "Virtual Environments and Advanced Interface Design", New York, Oxford University Press, 1995.

26. Calhoun, G. L. and Janson, W. P., "Eye Line-of-Sight Control Compared to Manual Selection of Discrete Switches", US Air Force report AL-TR-1991-0015, NTIS: AD-A273 019, 1991.

27. Borah, J., "Investigation of Eye and Head Controlled Cursor Positioning Techniques", US Air Force report AL/CF-SR-1995-0018, September 1995.

28. Calhoun, G. L., Janson, W. P. and Arbak, C. J., "Use of eye control to select switches", in "Proceedings of the Human Factors Society - 30th Annual Meeting", 1986, pp 154-158.

29. Calhoun, G. L. and Janson, W. P., "Eye control interface considerations for aircrew station design", in "Sixth European Conference on Eye Movements", Leuven, Belgium, 1991.

30. Ware, C. and Mikaelian, H. T., "An evaluation of an eye tracker as a device for computer input", in by Carroll, J.M. and Tanner, P. P., (Eds) "Proceedings of Human Factors in Computing Systems and Graphics Interface Conference", Toronto, Canada. 1987, pp. 183-188.

31. Hatfield, F., Jenkins, E. and Jennings, M. W., "Eye/Voice Mission Planning Interface (EVMPI)", US Air Force report TR-J103-1, 1995.

32. Glenn, F. A. Harrington, N., Iavecchia, H. P., and Stokes, J., "An Oculometer and Automated Speech Interface System", in "Analytics, Technical Report 1920", Analytics, Willow Grove, PA, May 1984.

33. Bolt, R. A., "The Human Interface: Where People and Computers Meet", Lifetime Learning Publications, London, UK. 1984.

34. Koons, D. B., Sparrell, C. J. and Thorisson, K. R., "Integrating Simultaneous Input from Speech, Gaze, and Hand Gestures," in Maybury, M. T. (Ed.) "Intelligent Multimedia Interfaces", Menlo Park, AAI Press/The MIT Press, 1993.

THE TECHNOLOGY AND APPLICATIONS OF GESTURE-BASED CONTROL

Grant R. McMillan
Air Force Research Laboratory (AFRL/HECP)
2255 H Street
Wright-Patterson Air Force Base, OH, USA 45433-7022

SUMMARY

This lecture reviews the technology for using hand, body and facial gestures as a means for interacting with computers and other physical devices. It discusses the rationale for gesture-based control technology, methods for acquiring and processing such signals from human operators, applications of these control technologies, and anticipated future developments.

1. INTRODUCTION

"Body Language" is an important component of normal interpersonal communication. Gesture-based control seeks to exploit this channel for human-machine interaction. Because traditional input devices constrain the expressive power of the human hand, scientists and engineers are developing a variety of techniques to read hand and body movements directly. For example, most currently available interfaces only make use of discrete pieces of data produced by the user's movements. This sometimes stems from the use of intrinsically discrete input devices, such as a keyboard or numeric keypad. Even with continuous input devices, such as mice, only specific events and data points (e.g., the co-ordinates of the pointer when the user clicks) are taken into account by most applications. Gesture-based interaction attempts to take advantage of the continuity and dynamics of the user's movements, instead of only drawing discrete information from these movements.

Although the terms are sometimes used loosely, **gesture** formally refers to dynamic hand or body signs, while **posture** refers to static positions or poses.

2. THE RATIONALE FOR GESTURE-BASED CONTROL

Gesture is a very natural human communication capability. Therefore, it should lend itself to easily learned interaction techniques. A distinguishing feature of the gesture communication channel is that it allows one to act on one's environment as well as to retrieve information from it. Three complementary and interdependent functions of gesture are pointed out by Cadoz [1]:

- The **epistemic** function, which corresponds to perception. This includes:
  - the haptic sense, which combines tactile (touch) and kinaesthetic sensations (awareness of the position of the body and limbs), and gives information about size, shape and orientation.
  - the proprioceptive sense, which provides information on weight and movement through joint sensors.
- The **ergotic** function, which corresponds to actions applied to objects.
- The **semiotic** function, which concerns communication. Examples include sign language and gesture accompanying speech.

In this lecture we are primarily concerned with action and expression, thus the ergotic and semiotic functions. However, feedback through the epistemic function is important in some gesture applications.

Typical gesture commands are terse and powerful. A single gesture can encompass a command as well as its arguments. For example, one gesture can combine the point and click operations of a mouse. Taking into account the user's movements, in all their continuity and dynamics, can provide more information than current interfaces and enrich the interaction. For instance, in a drawing program, a linear trajectory can be interpreted as a line-drawing command, while a curved trajectory would start the drawing of a circle. More abstractly, a cross drawn on an object can be a command for deletion; this would be an iconic use of gesture. Even further, provided that adequate tracking devices are used, three-dimensional trajectories and the postures of the limbs can be considered, allowing gestures to be recognised more precisely and making direct gestural interaction possible. This can provide for more natural control of a system at a lower cognitive cost. As a matter of fact, the hand can become the actual input device being used.

The preceding considerations apply, primarily, to intentional gestures. Some gestures, such as lip movements during speech, are not generally deliberate and typically provide contextual information or are interpreted jointly with another means of communication. Other spontaneous gestures accompanying speech do not constitute a language, but work has been done on typologies, e.g., gestures can stress specific words or sentences, indicate an object or place (deictic gestures), or sketch a shape or picture.

3. CHARACTERISTICS OF HUMAN GESTURES

The body movements involved in gestural communication can be a source of fatigue; thus it is important to use concise and simple-to-execute gestures. High precision cannot be relied on over time, and as is the case with gaze, it is very difficult for a human to maintain a static posture.

While the kinaesthetic sense gives one an indication of the position of the body and limbs, it is not sufficient to ensure that the desired gesture was adequately produced. Hence, feedback on gesture recognition is required.

Gesture is made more difficult in a dynamic environment. As with head-based control, hand movements are impaired by G forces and by vibration. So [2] investigated the transmission of vertical seat vibration to the outstretched hand at frequencies up to 10 Hz, and found involuntary hand motion in both the vertical (pitch) and lateral (yaw) directions. The vertical disturbance produced a resonant peak for the hand at about 2 Hz. Amplitude of hand motion in the lateral axis rose

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

gradually to about 5 Hz, beyond which it had a fairly flat response.

Gesture is characterised by large intra- and inter-subject variability. The difficulty of precisely reproducing a gesture is a potential source of precision and recognition problems. Differences between individuals suggest that some training of the recognition system is generally needed.

Another problem in free gesture recognition is similar to one encountered in natural speech understanding. A continuous stream of position data is received and has to be converted into a series of gestures considered as lexical entities. A further complication is the fact that co-articulation of gestures modifies the individual gestures, as is the case with phonemes. This leads to the problem of defining and recognising the beginning and ending points of a gesture. A number of systems avoid these difficulties entirely by limiting recognition to static postures.

Still another issue to be dealt with is the "immersion problem", especially in the case of unobtrusive methods of gesture capture. If every movement is subject to interpretation by the system, the user will be deprived of interpersonal communication for fear that a movement could be acted upon by mistake. The only solution is to provide an effective and unobtrusive way of detecting whether a gesture is addressed to the recognition system.

4. THE TECHNOLOGY FOR ACQUIRING AND PROCESSING GESTURE COMMANDS

Human gesture can be captured using a variety of hardware devices. Contact devices, besides classical ones such as mice, trackballs, trackpads and touch screens, include a variety of more exotic items such as spaceballs, 3-D mice and so on. The head, hands and body can be localised in space using trackers, video techniques, gloves or suits. Trackers are devices that allow one to directly measure the position and orientation of a body part in space. Video techniques use image recognition in order to follow a specific body part and then reconstruct its position, orientation and posture from 2-D video images. Gloves and suits allow one to measure the relative positions and angles of body components. A comprehensive directory of manufacturers of input technologies, of which a large part is devoted to gesture capture devices, is available in [3].

Among the criteria to be taken into account when evaluating gesture capture devices are the following:

- Accuracy - expected measurement error
- Range - an area or volume in which measurements can be made (accuracy is often specified for a given range)
- Precision - the repeatability of measurements
- Resolution - the smallest measurable physical change
- Update rate - the measurement frequency
- Latency - the time the system takes to report a physical change
- Cost
- Dependability

4.1. CONTACT DEVICES

These devices work in two dimensions and usually operate through a straightforward translation into two-dimensional screen space. Such devices can, however, be used for more advanced purposes, such as two-dimensional gesture or handwriting recognition. Some of these devices, e.g., graphic tablets, allow one to use contact to signal the beginning and ending of gestures, which addresses the segmentation problem noted above. The most serious limitation of these devices is that gesture is quite constrained. In particular, one hand is generally completely involved with the contact device.

Two kinds of contact devices can be distinguished: direct ones, which allow one to point on the screen surface, and indirect ones, for which interaction is mediated by a translation into screen space. Indirect pointing devices require additional co-ordination in that the operator has to match his or her movements with displacements in a different plane.

Direct pointing devices include the following:

- Lightpens are attractive but must be picked up, lead to arm fatigue (if the screen is vertical), and obstruction of the screen by the hand.
- Touchscreens (capacitive, ultrasonic, resistive or using a matrix of light beams) are fairly easy to use and robust. Although they allow good precision, since the fingertip is very sensitive and accurate, it is nearly impossible to be precise on first contact. One way to achieve this aim would be to have the screen "sense" the finger as it approaches, provide appropriate feedback and perform an action only at contact. This approach has been tested on prototypes, but is not yet an available technology. Other prototype touchscreens allow actions involving more than one finger and even sense forces tangential to the screen surface. Touchscreens have the same arm fatigue and screen obstruction problems as lightpens and produce the additional problem of screen smudging.
- Styluses (used in some notebook computers) are more comfortable and precise. They allow handwriting, but must be picked up and arm fatigue problems can occur if they are not used on small screens where the hand has a resting point.

Indirect pointing devices include the following:

- The mouse (optical, physical or acoustic) is very precise and rapid, but must be grasped and requires some desk space. Movement can be hampered by the wire, except for modern infrared-equipped models.
- Trackballs have the same use as mice but occupy less desk space.
- Joysticks are fast and efficient for direction changes and small movements. They are good for tracking targets. Some force feedback is possible. Absolute joysticks map the position of the pointer to the position of the stick, while isometric or velocity-controlled joysticks map pressure on the stick to velocity of the pointer. An example of the latter is the finger-operated mouse replacement found on some portable computers.
- Graphics tablets (resistive, magnetic or acoustic) offer good performance for writing or drawing. Modern models are sensitive to stylus pressure, allowing for very elaborate forms of expression. These devices are comfortable and precise but often require significant desk space.
- Touchpads present the same advantages as touchscreens, without obscuring the screen. Some training is required to establish co-ordination.

A more detailed summary of contact devices can be found in [4, 5].

4.2. TRACKERS

This section provides an overview of various devices that allow one to measure, in real time, the position of an object in space, that is, the six parameters (three co-ordinates and three angles) that correspond to its six degrees of freedom. These devices can be used for tracking the head, hands or other body components. They have also been used for person localisation and body posture recognition, although in the latter case the number of devices affixed to the body can make such systems awkward. Tracking can be done using mechanical connections to potentiometers or non-contact techniques such as magnetic fields, ultrasonic or infrared beams, or radar.

4.2.1. Mechanical Tracking

Mechanical tracking involves connecting the tracked object to its environment, using potentiometers linked to the object via articulated rods or cables. This allows very high update rates and very low latency. It is also an inexpensive solution. On the other hand, the usable range is small and the apparatus impairs free movement; the attachments to the body preclude most in-flight use and make these systems difficult to accept. This type of tracker has mostly been used for measuring head orientation.

4.2.2. Electromagnetic Tracking

Electromagnetic trackers include a transmitter, which is made up of three coils radiating orthogonal electromagnetic fields in a radius of a few meters. The mobile receiver element is also made up of three coils. It receives varying signals depending on its position relative to the transmitter. An electronic unit ensures proper modulation of the radiated fields, measurement of the currents in the receiver coils, filtering of the data and computation of receiver position. Some of these devices allow simultaneous measurement of the position of several receiver units.

These trackers are moderately expensive, but on the whole they are the most precise among the non-contact techniques. They also offer a large operating range. The main disadvantage of electromagnetic trackers is that any metallic object in the vicinity will generate induced electromagnetic fields and hamper measurements. Any source of electromagnetic radiation, such as a video monitor, can introduce errors as well. Also, there must be an electrical connection between the receiver and the electronic unit. This can limit free movement.

4.2.3. Ultrasonic Tracking

These trackers make use of ultrasonic pulses to compute distances based on time propagation measurements. The main advantage of these trackers is that they work seamlessly in metallic environments. They also tend to be less expensive than electromagnetic ones. However, ultrasonic trackers face a directivity problem. Receiver units must have direct line of sight to the emitter. The latency is greater than with other trackers since it includes the propagation of ultrasonic waves. Furthermore, since the speed of sound varies with temperature, temperature variations lead to errors. Other limits stem from the compromise that must be made in the choice of frequency. Too high a frequency will decrease the range since air attenuates ultrasonic waves. Useful range at 80 kHz is limited to about 2 meters. The usable range will be decreased further since directivity increases with frequency. With low frequencies precision is limited by wavelength (4 mm at 80 kHz). Some new trackers continually measure phase shift between the source and receiver, which leads to improvements in precision and latency. Other problems with this tracking technology are its sensitivity to ambient sound perturbations, as well as to reflections off walls.

4.2.4. Optical Tracking

Optical trackers generally use infrared light emitting diodes (LEDs). Most of them are built for specific needs. They can be divided into those that use point receivers, e.g., phototransistors, and those that make use of planar receivers, such as cameras.

Planar devices measure the location of a point light source (or a reflective marker) using multiple cameras. The 2-D marker positions detected by each camera are correlated to compute the 3-D co-ordinates. Using markers relieves the need for an attached wire to provide power, but makes image processing more difficult unless external illumination is provided. The cameras can either be fixed and track mobile markers (outside in), or be mobile and track fixed beacons (inside out). The outside in approach limits precision. The cameras must have a wide field of view, yet measure small movements. The inside-out approach provides better results if there are enough beacons in the environment. However, camera size and weight can be a problem.

These devices face the same directivity problem as ultrasonic trackers, and the usable range is similarly limited. Furthermore, they can be perturbed by light and the use of infrared light makes them impossible to use in combination with night-vision goggles. In a military context, possible remote detection of infrared sources can be a cause for concern.

4.2.5. Other Trackers

Non-contact, electric field sensing techniques are under development which enable 3-D position tracking without encumbering sensors or cables [6]. Movements of a body segment immersed in a dipole field are sensed as changes in displacement current to ground. While these systems can track the position of a large body segment, such as a hand, they do not yet have the resolution to track individual fingers.

Recently, a variety of low cost emitter-less trackers have begun to appear using principles similar to aircraft and missile inertial guidance systems. The sensing components may include inclinometers, Hall-effect compasses, gyroscopes or accelerometers. Inclinometers measure orientation with respect to gravity and are sensitive to other sources of acceleration. Compasses find the north magnetic pole and are perturbed by magnetic fields and metallic masses. Gyroscopes either use rotating masses or piezo-electric crystals. They only permit relative measurement, as do accelerometers. This leads to integration error accumulation when absolute position must be computed. Despite these constraints, adequate performance can be achieved for many applications.

Table I provides a summary of the different tracking technologies.

4.3. COMPUTATIONAL VISION SYSTEMS

These systems use classical image recognition techniques to find silhouettes of the hands or body and, in-turn, to identify postures. Figure 1 shows the layout of a system used for cursor control based upon finger pointing [7]. The computational problems are even more challenging than with marker-based optical systems, if real-time operation is required. Limited camera resolution necessitates a compromise between adequate recognition of small elements (such as fingers) and the large field of view necessary for free movement. Obstruction of the fingers by the hand or other body segments is another problem. Correlating several sources in order to compute 3-D information, though a workable solution for simple gestures

Table I. Comparison of Trackers

Type of Tracker	Range	Precision	Cost	Comments
Mechanical	Limited	Very good	Low	Bulky, constrains free movement
Electromagnetic	Large	Good	Moderate	Sensitive to magnetic fields and metal objects
Ultrasonic	Visible area	Moderate	Low	Sensitive to temperature, humidity and sound
Optical	Visible area	Good	Variable	Sensitive to light
Inclinometers, compasses	Unlimited	Moderate	Low	No position measurement
Gyroscopes, accelerometers	Unlimited	Low for position (integ. Errors)	Low	Shock-sensitive

such as pointing [8], is far from trivial. A problem common to all video techniques is that even 60 frames per second, the current limit for typical video cameras, is not sufficient to follow rapid hand movements.

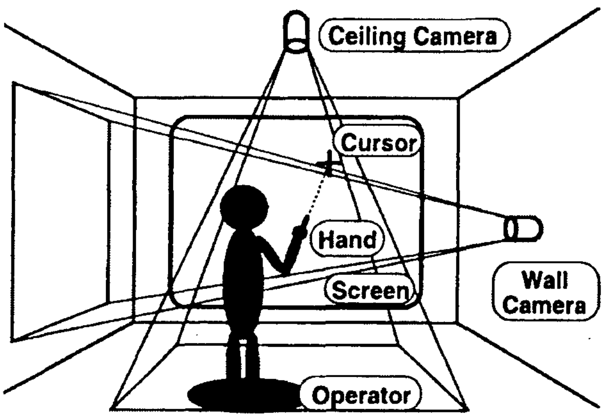


Figure 1. Camera and screen layout for video-based detection of finger-pointing direction. Finger direction, in-turn, controls the position of a cursor on a large screen display. From [7].

4.4. GLOVES

Gloves measure hand and finger angles and movements of the fingers relative to the hand. Most can be equipped with a position tracker in order to follow global hand position. Numerous sensors are needed and the resulting data rate can be high. Various measurement technologies can be used, including optic fibres, Hall effect, resistance variation or accelerometers. Gloves have been used as pointing devices, but they offer a much richer form of interaction through hand posture recognition and dynamic gesture interpretation. The main problems encountered are repeatability, precision and reliability. Almost every glove needs calibration before each use, since the manner in which it is fitted onto the user's hand greatly affects the measurements. Sensor technologies used in gloves have been applied to body posture recognition using "data suits", but this field is still fairly immature.

The first widely known glove, the DataGlove, appeared on the market in 1987 (Figure 2). It takes advantage of the attenuation of light in bent optic fibres to compute joint flexion. It uses ten sensors (two on the lower joints of each finger, and two on the thumb) and works at 60 Hz. Its accuracy is on the order of 5-10 degrees; it is limited because attenuation is not a linear function of joint angle. This precision is insufficient for complex gesture recognition. Another drawback of this technology is that light attenuation becomes permanent after repeated use and the fibres must be replaced. The fibres are also fairly fragile. Production of this glove has been discontinued, but General Reality Company is selling a model based on fibre optic technology.

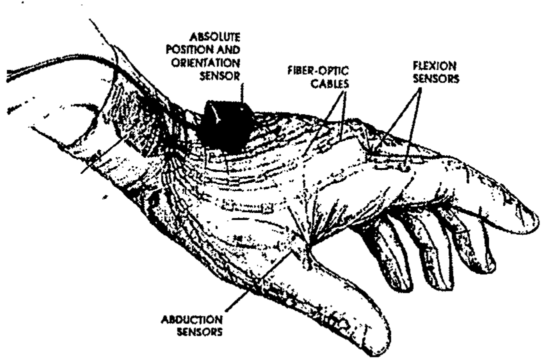


Figure 2. Schematic of DataGlove with magnetic receiver attached for tracking hand position and orientation.

Game designer Nintendo introduced the Powerglove in 1989 as a game controller. It is a very inexpensive device that uses the variation in conductivity of carbon ink tracks to measure flexion. It is coupled with a low-cost ultrasonic tracker. Production has been stopped, not due to the modest performance, but because the dedicated game market was not well enough developed.

A much more sophisticated device, the CyberGlove, employs 18 or 22 foil strain gauges for measuring flexion. Two are used for thumb joints, two or three for finger joints, four for abduction (thumb, middle-index, middle-ring and ring-pinkie), two for palm arch (thumb and pinkie) and two on the wrist (pitch and yaw). The operating rate can be as high as 149 Hz and the accuracy is about 1 degree. This glove is quite expensive but provides very good performance.

A considerably more elaborate model is the Dexterous HandMaster (Figure 4). It includes an exoskeleton with Hall effect sensors (up to four per finger) located in each joint. It can measure joint angles with a frequency of up to 200 Hz. Sensitivity and resolution are high, but calibration problems remain. The glove is also fairly heavy (350 grams).

The SensorGlove [9] is a more recent experimental device, which uses accelerometers. It does not allow accurate position measurement (the required double integration leads to error accumulation), but is usable for dynamic gesture recognition.

Accelerometers allow excellent update rates (up to 5 kHz) and are lightweight devices, but they are sensitive to shock. Also, it is not clear how they would behave in high-acceleration environments such as a cockpit.

Table II summarises the essential aspects of different glove technologies.

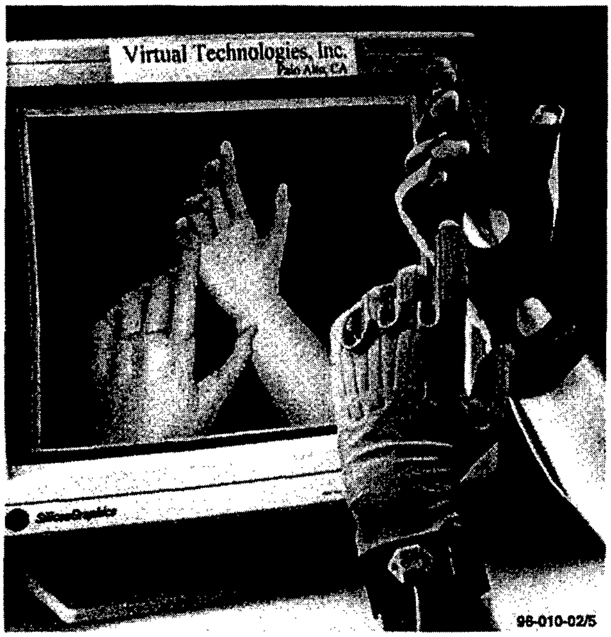


Figure 3. Photo of the 18-sensor CyberGlove and VirtualHand software display (Courtesy of Virtual Technologies Inc., Palo Alto, California).

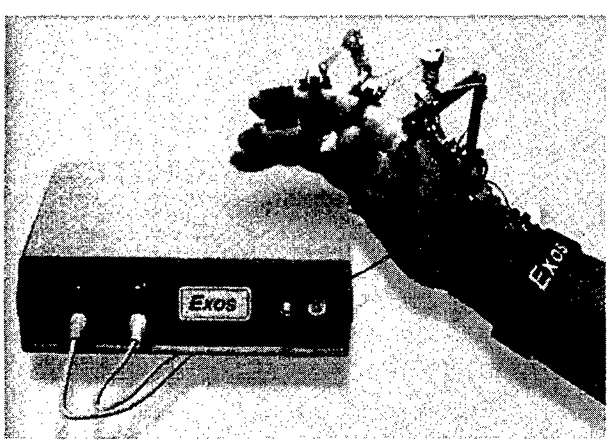


Figure 4. Photo of the Dexterous HandMaster (Courtesy of Exos, Inc., Woburn, Massachusetts).

4.5. OTHER DEVICES

3-D mice are devices designed to control a pointer in a three-dimensional space. They can use the same technologies as trackers, but are typically designed as generalisations of desktop mice. Typical 3-D mice can be moved in a spherical radius of about 25 cm, with a precision of less than 1 mm and a measurement frequency of 250 Hz.

Spaceballs are spheres that allow one to control six degrees of freedom. They sense force applied on each axis and torque around each axis. They are rather inexpensive devices. The main drawback is whether the user can attain real independence between these six degrees of freedom. For example, it is difficult to apply a linear force with no torque at all.

In order to widen the usable range of gesture recognition, "smart ceilings" [10] and "smart floors" are being investigated. Smart ceilings use a network of LEDs and head-mounted photo-receivers and allow one to detect the position and orientation of an operator in a room. Smart floors consist of a matrix of pressure sensors that give information on the position of the user, but could also determine movement direction and speed, and might assist with user identification since they can estimate weight.

4.6. SIGNAL PROCESSING AND CONTROL ALGORITHMS

4.6.1. Hand and Body Gestures

The algorithms for determining positions and joint angles, based on the sensor inputs, are provided with the systems described above. With the glove-based systems, some individual user calibration is required. Magnetic trackers require the mapping of ferromagnetic and metal conductive surfaces in the user's environment, but no individual user calibration. For interactive applications that employ rapid body movements, one may need to add or modify movement prediction algorithms to compensate for system delays. Although general-purpose posture recognition software is becoming available for the glove-based systems, the development of algorithms to recognise specific postures or gestures is often left up to the user. These algorithms tend to be application specific, although some general approaches are common. For example, the recognition of a fixed set of hand postures is often based on look-up tables that contain acceptable ranges of values for each position and joint measurement.

Interpreting gestures is a much more challenging problem since pattern analysis must be performed on a moving hand. Many approaches compare the motion vectors for each degree-of-freedom of the hand to reference vectors representing the target gesture. This match must be within error tolerances and these tolerances are weighted by the contribution of each hand motion vector to gesture discrimination. The weighting may be accomplished with principal components analysis [11], with Bayesian rule-based techniques [12], hidden Markov models [13], edge-based techniques [14], a "sum of squares" method [15], or it may be performed by a neural network [16]. A different technique that has shown promise in several applications is a feature analysis approach developed by Rubine [17]. Originally developed for the interpretation of 2-D written gestures, it has been extended to 3-D hand gesture recognition [18]. The features Rubine analysed were pre-specified measurements of the 2-D movement trajectories, such as sines and cosines of the initial angle of the gesture, the duration of the gesture, and so on.

For telemanipulation and robot control applications resolution of the kinematic differences between the human hand and the robot hand is often required. Algebraic transformations have been employed to perform this human-to-robot mapping. Alternatively, the kinematic differences can be resolved by determining the 3-D position of the user's fingertips and driving the robot's fingertips to match.

Table II. Comparison of Glove Technologies

Glove Technology	Precision	Cost	Comments
Optic Fibre	Low	Low	Fragile, subject to wear
Strain Gauges	High	High	
Resistive Ink	Very low	Very low	
Hall Effect	High	Very high	Cumbersome
Accelerometers	Low for position measurement	Prototype only	Sensitive to acceleration and shock

Segmentation is a difficult challenge with dynamic gesture recognition. As is the case with continuous speech recognition, co-articulated gestures interfere with the detection of individual gestures. It is also a nontrivial problem to identify the beginning and end points of a gesture. Typical solutions require the operator to take a "default" hand posture between gestures which serves as an anchor for the system. Davis and Shah [19] demonstrate the feasibility of using simple finite state machines under this paradigm.

4.6.2. Facial Gestures

The human face supports a variety of communicative functions, such as identification, perception of emotional expressions and lip-reading. Lip-reading is discussed in detail in the lectures on speech-based control and to some extent in the lecture on biopotential-based systems. Face perception is currently an active research area in the computer vision community. Much research has been directed towards feature recognition in human faces. Three techniques are commonly used for dealing with feature variations: correlation techniques, deformable patterns, and spatial image invariants.

Several systems for locating faces have been reported. By moving a window covering a subimage over the entire image, faces can be located within the image. Sung and Poggio [20] report a face detection system based on clustering techniques. The system passes a small window over all portions of the image, and determines whether a face exists in each window. A similar system with better results has been reported by Rowley et al. [21]. A different approach for locating and tracking faces is described in Hunke and Waibel [22]. This system locates faces by searching for skin colours in the image. After locating the face, the system extracts additional features to match a particular face.

Another active research and development area is the recognition of facial expressions. This work combines techniques for tracking and locating the face with the recognition of different expressions such as disgust, anger, happiness and surprise. The goal is to develop an intelligent interface that would adapt to the user based on the emotional state determined from his or her facial expressions. Examples of this work are presented by Essa and Pentland [23] and Yacoob and Davis [24].

4.6.3. Control Modes or Styles

Given that a specific hand gesture, posture or facial expression can be reliably discriminated from other activity, the dialogue designer still must determine how it will be used for interaction with a system. Sturman and Zeltzer [25] describe several modes of control that encompass many of the available dialogue options. First, the designer can choose to use discrete or continuous features of a gesture in the dialogue. Within each of these categories there are three styles of input:

- Direct - Features of the gesture or posture generate kinematically similar actions in the task domain. One-to-one control of a robot hand would be a good example of this style of interaction.
- Mapped - Features of the gesture or posture are mapped in some logical fashion to actions in the task domain, but there may be no kinematic similarity between the features and actions. For example, the number of raised fingers might indicate which of four levels of force should be applied, or circling of the index finger might indicate that a load on a crane is to be lifted.
- Symbolic - Features of the gesture or posture are interpreted as commands to the system. While this may be similar to the mapped style, it may also be significantly more abstract and may employ knowledge-based reasoning to determine the intention or emotional state of the operator. Most interpretations and uses of facial expressions would fall into this category.

5. USER FEEDBACK REQUIREMENTS

In many applications the only feedback that is provided or required is the system's response to the recognised gesture. Examples include simulated movement in the direction toward which the user is pointing and synthesised speech following recognition of a sign language gesture. Feedback requirements for applications that involve simulated object manipulation, vehicle control and robot operations are still the subject of research and development. In each of these cases tactile and kinaesthetic feedback play an important role in normal human-system interaction. These cues are absent in most gesture-based systems. Significant progress is being made in the development of force-reflection [26] and tactile stimulation systems [27] that can provide this feedback through normal sensory modalities. In addition, there is evidence that substitute feedback can be provided with vibrotactile, auditory and electrotactile displays [28]. However, tactile and kinaesthetic feedback are not required in all cases. Massimino and Sheridan did not find enhanced performance of a peg insertion task when artificial or actual force cues were provided. As described in the lecture on biopotential-based control, users of EMG-controlled prosthetic arms can perform a grip force control task adequately with visual feedback alone, although performance is slightly enhanced when synthetic pressure cues are provided. The importance of simulated tactile and kinaesthetic feedback depends on the specific task, the experience of the user, the availability of substitute visual and auditory cues, and the implementation of the artificial feedback. Until additional parametric studies are performed, it is difficult to provide specific guidelines. Nevertheless, it seems clear that some form of tactile and kinaesthetic feedback

will be required for certain object manipulation and tool operation tasks in many telerobotic applications.

Rather than attempting to simulate the sensations that would be present in object manipulation or vehicle control, gesture feedback can be used in a more abstract fashion, such as the following example [29]. As reported here, the operator draws on a map and the drawing tool produces a force feedback proportional to the population density gradient. This immediately allows the user to determine, for example, the least disruptive highway route by simply following the path of least resistance.

A distinction must be made between tactile and force feedback. Tactile feedback provides information on the nature of the surface of a grasped object (geometry, roughness, temperature) while force involves the proprioceptive sense and provides information on the elasticity, weight and movement of an object.

Evaluation criteria for feedback systems include bandwidth (this determines, for example, the quality of a simulated texture) and available power of a force feedback system. There is a compromise to be reached here, since great forces are needed to simulate a hard object, but misapplied forces could be harmful to the user. Feedback delay is also an important factor. Delayed force feedback is useless and can even make a system unusable.

5.1. TACTILE FEEDBACK

Pneumatic, shape-memory materials and vibrotactile technologies have been used for providing tactile feedback. Experiments have also been performed using hydraulic systems, electric stimulation of the skin or even direct neuro-muscular stimulation. The currently available devices are few and this area is still mostly a research domain.

Pneumatic devices use a number of small balloons, generally integrated into a glove, which can be inflated to apply pressure on the fingers or palm. A matrix of micro-rods, made of a shape-memory material, has been used for tactile stimulation. The rods change shape when heat is applied and are suitable for miniature devices. Vibrotactile devices use small loudspeakers, and electromagnetic or piezo-electric micro-rods, which transmit audio frequency (around 200 Hz) vibrations to the skin. They are most appropriate to simulate texture of a virtual object. Some experiments have added thermal stimulation in order to indicate emergency conditions, for example.

5.2. FORCE FEEDBACK

Force feedback systems can use electric, hydraulic or pneumatic technologies. They were first applied to telemanipulation arms. Increasing miniaturisation has allowed the incorporation of such systems into gloves and joysticks. Despite this progress, the main disadvantage is that most of these systems remain bulky and rather intrusive, which prevents their use in transportable or wearable devices.

6. APPLICATION EXAMPLES

6.1. TELEOPERATION AND ROBOT CONTROL

Remote manipulation of objects because of weight or exposure risks, e.g., radioactivity, has been performed for many years using direct mechanical linkages or electric motors that permit force amplification. Even though these systems do not actually include a computing system, they involve the transmission of gestural information. In that respect they are forerunners of a number of object manipulation applications.

Hale [30] used a DataGlove to control a robot arm in a task that required retraction, slewing and insertion of a block in a test panel. He compared his results to another study that used a conventional six degree-of-freedom handcontroller as the input device. He concluded that performance with the DataGlove compared favourably with the "standard" device and that it provided a natural and intuitive user interface. Brooks [31], on the other hand, was less optimistic about the DataGlove for robot control; his evaluation involved more complex gestures and a neural network for gesture recognition. The reader should recall, however, that the DataGlove is very limited for precise manipulation tasks.

Sturman and Zeltzer [25] evaluated gestural control of a six-legged mobile robot with manipulator arms. They compared whole-hand input using a DataGlove to conventional input using a set of dials. Three different levels of control were investigated. For the control of low-level walking, the whole-hand interface was superior, since it took advantage of natural co-ordination patterns when one produces walking motions with their fingers. For object manipulation, the two controllers were roughly equivalent. For high-level steering, the whole-hand interface was inferior, because of hand instability and the difficulty of exercising control at extreme rotations of the wrist.

6.2. VIRTUAL AND AUGMENTED REALITY

Several examples are provided by 2-D and 3-D displays in which the user can touch, grab and move objects by pantomiming these activities with glove-based sensors. In these applications the user actually sees a computer rendering of their hand performing the object manipulations. Researchers at NASA/Ames have used this approach in a virtual wind tunnel to explore simulations of computational fluid dynamics. Aeronautical engineers can put their hands and head into a simulated fluid flow and manipulate the patterns in real time [32].

The GROPE project at the University of North Carolina [33, 34] is among the first applications to use force feedback for interacting with a computer simulation. The application domain is the simulation and graphical representation of interactions between complex molecules. A specifically-developed force feedback manipulating rod, allowing six degree-of-freedom movement is employed. When the user modifies the simulated position of one molecule by moving the rod, the simulation computes intermolecular forces and reflects them back through the feedback system. As a result of the computational time needed for the simulation, the system produces relatively low fidelity sensations. Nevertheless, this system allows one to begin to explore possible chemical bonds between molecules.

If an application requires the user to manipulate virtual objects in some way, accuracy of depth perception becomes an issue, particularly for computer-generated displays that lack the rich textural cues available in real life. Takemura, Tomono and Kobayashi [35], using a stereoscopic projector to display targets in 3-D space, found that subjects could "touch" the objects with a three-dimensional tracker with satisfactory accuracy. Ineson and Parker [36], using a similar task but with a head-mounted display, found that some subjects could "touch" the virtual objects with good accuracy, while others had great difficulty in judging their depth.

Augmented reality is often a more pragmatic approach. It consists of adding virtual elements to physical objects that one interacts with in the real world. It aims at integrating computing systems into the real world instead of embedding the user in a simulated world. An example is the Digital Desk demonstration [37] that allows one to work with paper and to

use a variety of digital tools at the same time. If the user is drawing, for example, a video camera can scan the drawing. It can then be digitally edited by means of a projector. It is even possible to mix both media and work on a partly-real/partly-electronic hybrid document.

Some virtual cockpit applications work in this fashion by projecting synthetic imagery onto the physical environment of the pilot. White et al. [38] were interested in the problem of interacting with real cockpit instruments when direct vision of the instruments was obscured. They set up a virtual keypad on a head-mounted display that overlaid a real keypad, and operated it using a finger-tracker.

6.3. SIGN LANGUAGE INTERPRETATION

Sign language interpretation continues to be a significant area for gesture research and development. This type of application is not the focus of this lecture and we will touch on it only briefly. Fels and Hinton [16] developed a hand gesture to speech system using a neural network. Their system mapped hand postures to complete root words, followed by a directional hand movement that modified the word ending (singular, plural, etc.) and controlled speech rate and emphasis. Performance of a single "speaker" with a vocabulary of 203 words was evaluated following a network training phase. With near real-time speech output, the wrong word was produced less than 1 percent of the time and no word was generated approximately 5 percent of the time. Similar hand gesture to speech demonstrations were conducted by Kramer and Liefer [12] with American Sign Language and by Takahashi and Kishino [11] with the Japanese kana manual alphabet. Several papers on the subject appear in [39]. Some currently available systems such as the CyberGlove have software to convert fingerspelled words from American Sign Language into synthesised speech.

6.4. COCKPIT APPLICATIONS

Ineson, Parker and Evans [40] compared a video-based finger tracker with several other designation mechanisms to select buttons on a virtual, head-down panel during simulated low-level flight. Feedback for contact with the button was a colour change. Activation of the button required depressing a switch on the Hands-On Throttle and Stick (HOTAS) for confirmation. The finger-tracker was poorly rated by the subjects since it removed the hand from the flight controls for a substantial period of time. Some subjects found the device awkward to use since it was necessary to keep the finger in clear view of the tracking cameras. Although the normal means of operating a button is to reach out and press it, the task is essentially two dimensional. Methods such as head pointing and stick-top cursor controllers are suitable mechanisms also, and both were preferred to the finger tracker. Finger pointing direction would have been more suitable than finger position, since it could have been operated with the hand on, or near, the controls. Voice control was the overwhelmingly preferred selection technique for this task.

A series of experiments carried out at Wright-Patterson Air Force Base, Ohio, USA [41-44] required true 3-D selection of targets from a head-down, 3-D tactical map. In this case an electromagnetic tracker was strapped to the back of the hand, resulting in more robust and responsive tracking than the video-based technique used by Ineson et al. The tracking volume was remote from the actual map so that hand movements were actually made in a space close to the aircraft controls rather than within the volume of the map. This hand volume was reduced in scale so that hand movements were small compared to the size of the map. The volume was divided into four depth planes, so accurate depth control was not required. The hand tracker worked well and was, in general, faster and more accurate than a three-dimensional joystick. If the tracking volume was made too small, selection accuracy was impaired. Voice selection was also used in some of the experiments [41, 44]. Unless the targets were labelled it was difficult to define a suitable vocabulary and the method was slow compared to hand movement.

Reising et al. [41] and Solz et al. [42] also investigated two methods for simplifying object designation - contact cueing (colour change when the cursor was within the target volume) and proximity cueing (automatic selection of the target nearest to the cursor). The latter was found to be particularly helpful.

Not only must one chose the selection device to suit the task, but also one must consider the environment in which the device is to be used. A positional hand tracker might be the preferred device in a relatively benign environment, but might become unusable under the acceleration and vibration levels found in a fast jet or helicopter. A 3-D joystick would have the advantage of supporting the hand, but the space needed to integrate such a device must then be considered. The glove required by a gesture recognizer might be incompatible with safety equipment or might interfere with other tasks requiring finger-tip sensitivity. System lags that are tolerable in a controlled experiment might become problematic when the user has to attend to several tasks at once. Environmental and integration issues such as these must guide the choice of control devices for a specific cockpit task.

6.5. OTHER APPLICATIONS

One of the earliest examples of a multimodal interface involving gesture was the "Put-that-there" demonstration described by Bolt [45]. This demonstration combined hand pointing and speech recognition to permit natural interaction with objects on a large screen display. Pointing direction was sensed with a magnetic tracker attached to the hand. The interface responded to commands such as "Name that X" or "Put that there", where "that" referred to the object being pointed at and the action was defined by voice input. This demonstration provided a compelling example of integrated alternative control that allowed the user to directly manipulate task objects. No visible control devices were imposed between the user and his or her task.

CHARADE [18] is a system designed for gesture-based control of computer-aided presentations to an audience. Wearing a DataGlove, the speaker points at the screen, which constitutes an "active zone", and makes a short hand gesture corresponding to the required command. The gesture is then matched to an internal model consisting of a start position, hand and arm movement, and a stop position. This scheme prevents the "immersion syndrome" in that the speaker can keep using gestures when addressing the audience. It also alleviates the problem of gesture co-articulation, and the careful choice of start and end positions makes recognition easier. The choice of tense postures as start positions and relaxed ones as end positions is also a helpful for the recognition device. A variant of the Rubine algorithm [17] is used. Sixteen commands such as "next/previous page", "next/previous chapter", "table of contents", "mark this page" or "highlight area" are available. Recognition rates of 90 to 98% have been reached by trained users.

7. DESIGN METHODS AND PRINCIPLES

Sturman and Zeltzer [25] have proposed a method to assist users in designing and evaluating whole-hand input for specific

tasks. A flow diagram of their design process is shown in Figure 5. In the first stage the designer must answer questions such as: "Can existing hand signs be used to perform the task?", "Does the task require co-ordination of many degrees of freedom?" and "Should the absence of an intermediary control device improve performance?". If the answers to these questions support the use of whole-hand input, the designer then begins an analysis process that: (1) breaks the task down into primitives, (2) specifies the co-ordination, resolution, endurance and other requirements for each task component, (3) determines whether hand capabilities can meet these requirements and (4) identifies whole-hand input devices that provide the resolution, reliability and sampling rates required to meet the task specifications. After completing these steps, the prototyping and interface evaluation process can begin.

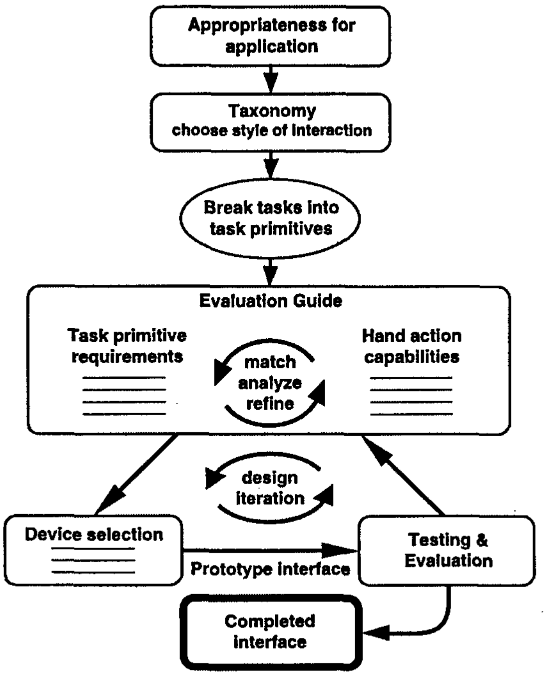


Figure 5. A method for designing and developing whole-hand input for specific applications and tasks. From [25].

Because of the relative immaturity of this area, detailed design principles are not available. Nevertheless, a review of the gesture literature suggests some general guidelines that are applicable to many situations:

- Gesture-based control should offer learning and performance advantages if the task is based on an already learned set of signs or signals. Glove-based translation of American Sign Language is an example.
- Gesture-based control should offer learning and performance advantages if the natural co-ordination of the body can be employed to co-ordinate multiple degrees of freedom in the external device. Finger walking to control the locomotion of a legged-robot is an example [25].
- Gesture-based control may be less effective than conventional control if the task requires high resolution control of a single degree of freedom. At least two factors contribute to this: (a) conventional controls often have higher resolution than gesture-based devices and (b) conventional controls often provide support and damping that is helpful in precision control situations. This may not be true for applications in which gesture affords more natural, user-scaled control location.
- Gesture-based control may be less effective than conventional control if tactile and kinaesthetic feedback is important for task performance.
- Gestures should be concise and quick in order to minimise fatigue. High precision over a long period of time should be avoided.
- Since most systems capture every motion of the user's hand, the controller must provide a well-defined means to detect the intention of gestures. An example is the CHARADE system [18] for controlling computer-based presentations to an audience. Gestures are acted on only when the user is gesturing within the "active zone" of the projection screen. Gestures to the audience are not recognised.

8. FUTURE DEVELOPMENTS

A significant disadvantage of existing gesture-capturing devices is that most of them limit the user's freedom of movement. This results from the need to grasp a sensor component, from wires attached to sensors or from limited sensor range. Progress in component miniaturisation and telemetry will help to solve this problem.

Static posture recognition has made great progress and allows reasonably high recognition rates, provided the user performs a standard procedure such as pointing at a target area or assuming a standard posture prior to issuing a command. This is not yet true for dynamic gesture recognition and software techniques are still developing in this field. The main difficulty is segmentation, i.e., detecting gesture beginning and end points. Aids such as hand speed and tension are currently being investigated.

General interface problems such as immersion are still not solved in a comprehensive fashion. The definition of an active zone partly solves this problem but may not be adequate for all applications. The development of adequate interface paradigms for gesture interaction with computers is still under active research; a consensus on the integration of gestures in interfaces is far from being reached.

In the feedback domain, the determination of appropriate stimuli is largely in its infancy. Appropriate modelling of physical objects and of their interaction with body parts is a prerequisite.

At the present time, learning how to operate a gesture-based interface is mostly done by example. Gesture notation is undergoing a significant amount of research. An example is HamNoSys (Hamburg Notation System) [46], which is a general iconic notation for sign languages. Although it initially aimed at notation of human sign languages, it has since proved helpful in the design of artificial gesture languages.

Despite these challenges, gesture-based applications are beginning to take advantage of the dexterity and natural co-ordination of the human body and to reduce the constraints of conventional input devices. In addition to the explicit control applications that have been the focus of this lecture, gesture also plays an important expressive role in human communication. While we use gestures to indicate specific actions and desires, we also use them to indicate emphasis and emotion. Interface designers are beginning to explore the recognition and application of facial expressions and other emotive inputs. Here deviceless, free-form gesture recognition

will be required, and an effective system will undoubtedly integrate the inputs from a variety of the alternative controls reviewed in this series of lectures.

9. ACKNOWLEDGEMENTS

This lecture draws most of its material from a chapter on Gesture-Based Control prepared by Dr. Pierre Dauchy, IMASSA-CERMA, France, as part of the final report of AGARD Working Group 25, entitled Alternative Control Technologies. The author wishes to thank Dr. Dauchy for his kind permission to use this material.

10. REFERENCES

1. Cadoz, C., "Le Geste Canal de Communication Homme/Machine, la Communication 'Instrumentale'", Technique et Science Informatiques, 13, 1, 1994, pp 31-61.

2. So, R.H.Y., "Comparison of the Transmission of Vertical Seat Vibration to the Head and Finger in a Stationary Target Aiming Task", United Kingdom and French Joint Meeting on Human Response to Vibration, 1988.

3. Buxton, B., "A Directory of Sources for Input Technologies", 1998, available on the Internet as http://www.dgp.utoronto.ca/people/BillBuxton/InputSources.html.

4. Dix, A., Finlay, J., Abowd, G., and Beale, R., "Human-Computer Interaction", UK, Prentice-Hall, (ISBN 0-13-437211-5), 1993, pp 59-63.

5. Shneiderman, B., "Designing the User Interface - Strategies for Effective Human-Computer Interaction", 2nd Edition, Addison-Wesley, 1992, (ISBN 0-201-57286-9), pp 244-253.

6. Zimmerman, T.G., Smith, J.R., Paradison, D.A., Allport, D., and Gershenfeld, N., "Proceedings of the CHI", 1995.

7. Fukumoto, M., Suenaga, Y., and Mase, K., "Finger-Pointer": Pointing Interface by Image Processing", Computers and Graphics, 18, 5, 1994, pp 633-642.

8. Fukumoto, M., Mase, K., and Suenaga, Y., "Real-Time Detection of Pointing Actions for a Glove-Free Interface", in "IAPR Workshop on Machine Vision Applications", December 1992, pp 473-476.

9. Hofmann, F., and Hommer, G., "Analyzing Human Gestural Motions using Acceleration Sensors", in "Progress in Gestural Interaction - Proceedings of Gesture Workshop '96", March 1996.

10. Ward, M., Azuma, R., Bennett, R., Gottschalk, S., and Fuchs, H., "A Demonstrated Optical Tracker with Scalable Work Area for Head-Mounted Display Systems", in "1992 Symposium on Interactive 3D Graphics", Association for Computing Machinery, 1992, pp 43-52.

11. Takahashi, T., and Kishino, F., "Hand Gesture Coding Based on Experiments using a Hand Gesture Interface Device", SIGCHI Bulletin, 23, 2, 1991, pp 67-74.

12. Kramer, J., and Liefer, L., "The Talking Glove: An Expressive and Receptive 'Verbal' Communication Aid for the Deaf, Deaf-Blind, and Nonvocal", Stanford, California, Stanford University Department of Electrical Engineering, 1989.

13. Starner, T., and Pentland, A., "Visual recognition of American Sign Language using Hidden Markov Models", in Bichsel, M., (Ed), "Proceedings of the International Workshop on Automatic Face- and Gesture- Recognition", 1995, pp. 189-194.

14. Uras, C., and Veri, A., "Hand Gesture Recognition from Edge Maps", in Bichsel, M., (Ed),"Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition", 1995, pp 116-121.

15. Newby, G.B., "Gesture Recognition using Statistical Similarity", in "Proceedings of the Virtual Reality and Persons with Disabilities Conference, 1993.

16. Fels, S.S., and Hinton, G.E., "Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer", IEEE Transactions on Neural Networks, 4, 1, 1993, pp 2-8.

17. Rubine, D., "Specifying Gestures by Example", Computer Graphics, 25, 4, 1991, pp 329-337.

18. Baudel, T., and Beaudoin-Lafon, M., "Charade: Remote Control of Objects using Free-Hand Gestures", "Communications of the Association for Computing Machinery", 36, 7, 1993, pp 28-35.

19. Davis, J., and Shah, M., "Gesture Recognition", University of Central Florida Technical Report CS-TR-93-11, 1993.

20. Sung, K. and Poggio, T., "Example-Based Learning for View-Based Human Face Detection", Massachusetts Institute of Technology AL Lab Technical Report 1521, 1994.

21. Rowley, H.A., Baluya, S., and Kanade, T., "Human Face Detection in Visual Scenes", Carnegie Mellon University Computer Science Department Technical Report CMU-CS-95-158, 1995.

22. Hunke, M., and Waibel, A., "Face Locating and Tracking for Human Computer Interaction", in "28th Asilomar Conference on Signals, Systems, and Computers", November 1994.

23. Essa, I.A. and Pentland, A.P., "Coding, Analysis, and Recognition of Facial Expressions", Massachusetts Institute of Technology Media Lab Technical Report No. 325, April 1995.

24. Yacoob, Y. and Davis, L.S., "Recognizing Facial Expression", University of Maryland Computer Vision Laboratory Technical Report CS-TR-3265, May 1994.

25. Sturman, D.J., and Zeltzer, D., "A Design Method for 'Whole-Hand' Human-Computer Interaction", ACM Transactions on Information Systems, 11, 3, 1993, pp 219-238.

26. Hannaford, B., and Venema, S., "Kinaesthetic Displays for Remote and Virtual Environments", in Barfield, W., and Furness, T.A., (Eds.), "Virtual Environments and Advanced Interface Design", New York, Oxford University Press, 1995.

27. Kaczmarek, K.A., and Bach-y-Rita, P., "Tactile Displays", in Barfield, W., and Furness, T.A., (Eds), "Virtual Environments and Advanced Interface Design", New York, Oxford University Press, 1995.

28. Massimino, M.J., and Sheridan, T.B., "Sensory Substitution for Force Feedback in Teleoperation", Presence, 2, 4, 1993, pp 344-352.

29. Negroponte, N., "Being Digital", Coronet, Hodder & Stoughton, 1995 (ISBN 0-340-64930-5).

30. Hale, J.P., "Anthropometric Teleoperation: Controlling Remote Manipulators with the DataGlove", NASA-TM-103588 (NTIS: N92-28521/2/XAB), 1992.

31. Brooks, M., "The DataGlove as a Man-Machine Interface for Robotics", in "2nd IARP Workshop on Medical and Healthcare Robotics", Newcastle upon Tyne, United Kingdom, 1989, pp. 213-225.

32. Bryson, S., and Levit, C., "The Virtual Wind Tunnel", IEEE Computer Graphics and Applications, 12, 4, 1992, pp 25-34.

33. Brooks, F.P., "Grasping Reality Through Illusion: Interactive Graphics Serving Science", in "Human Factors in Computing Systems - Proceedings of CHI '88", 1988, pp 1-11.

34. Brooks, F.P., Ouh-Young, M., Batter, J.J., and Kilpatrick, J., "Project GROPE: Haptic Displays for Scientific Visualization", Computer Graphics, 24, 4, August 1990.

35. Takemura, H., Tomono, A., and Kobayashi, Y., "A Study of Human-Computer Interaction via Stereoscopic Display", in Smith, M.J., and Salvendy, G., (Eds), "Work with Computers: Organizational, Management, Stress and Health Aspects", Elsevier Science Publishers, 1989, pp 496-503.

36. Ineson, J., and Parker, C.C., "The Accuracy of Virtual Touch", DRA Working Paper DRA/AS/MMI/WP95036/1, 1995.

37. Wellner, P., "Interacting with Paper on the DigitalDesk", Communications of the Association for Computing Machinery, 36, 7, July 1993, pp 87-96.

38. White, J.L., et al., "Virtual Cockpit Concepts: An Evaluation of Data Entry Techniques", DRA Report DRA/AS/MMI/CR95168, 1995.

39. "Progress in Gestural Interaction", in "Proceedings of Gesture Workshop '96", University of York, March 1996.

40. Ineson, J., Parker, C.C., and Evans, A., "A Comparison of Head-Out and Head-In Selection Mechanisms During Simulated Flight", DERA Customer Report DERA/AS/SID/510/CR97153, 1997.

41. Reising, J.M., Liggett, K.K., and Hartsock, D.C., "Exploring Techniques for Target Designation using 3-D Stereo Map Displays", International Journal of Aviation Psychology, 3, 3, pp 169-187.

42. Solz, T.J., et al., "3-D Stereo Displays: How to Move in the Third Dimension", SPIE Conference, Orlando, May 1995.

43. Solz, T.J., et al., "The Use of Aiding Techniques and Varying Depth Volumes to Designate Targets in 3-D Space", in "Proceedings of the 38th Annual Meeting of the Human Factors Society, 1994.

44. Reising, J.M., et al., "New Cockpit Technology: Unique Opportunities for the Pilot", SPIE Conference, Orlando, April 1994.

45. Bolt, R.A., "Put-That-There: Voice and Gesture at the Graphics Interface", Computer Graphics, 14, 3, 1980,pp 262-270.

46. Pritzwill, S., Leven, R., Zienert, H., Hanke, T., Henning, J., et al., "HamNoSys (version 2.0) Hamburg Notation System for Sign Languages / An Introductory Guide", International Studies on Sign Language and the Communication of the Deaf, 5, Hamburg, Germany, 1989.

APPLICATIONS OF SPEECH-BASED CONTROL

Timothy R. Anderson
Technical Advisor
Information Analysis and Exploitation Branch
Crew System Interface Division
Air Force Research Laboratory
Wright-Patterson AFB, OH 45433-7022

1. ABSTRACT

This lecture will examine many applications of speech based control in aerospace environments. Applications of speech recognition in fixed and rotary wing aircraft as well as in space and command and control will be discussed. Current performance of the technology and application problems will be presented. The lecture concludes with a discussion of required enhancements for aerospace applications.

2. AEROSPACE APPLICATIONS TO DATE

This section will examine current aerospace applications of speech-based control. Application and results in fixed wing and rotary wing aircraft as will as command and control and space will be presented. The use of speech as an indicator of operator state will be briefly discussed.

2.1. FIXED WING

One of the first series of flight trials of speech recognition equipment took place between 1982 and 1985, on-board a BAC 111 civil airliner. This particular aircraft was a flying laboratory, based at the Bedford, UK, airfield of the Defence Research Agency. A speaker-dependent connected speech recognizer, the Marconi SR128, was used to control the displays, radios, and the experimental flight management system. Average recognition accuracy was over 95% on a vocabulary that was built up over a period to about 240 words. Some pilots found the system so useful that they used it as a normal part of their cockpit interface, even during trials of other equipment. The cockpit environment of such an aircraft is, of course, much less noisy and stressful than that of most military aircraft.

The U.S. Air Force, NASA, and the U.S. Navy conducted a joint program in the mid-1980's to flight test interactive voice systems in the fighter aircraft. The program consisted of laboratory and simulator testing prior to flight tests. Significant improvements in recognition accuracy were made during each of the three phases of the program. Speaker-dependent, isolated-word speech recognition systems were evaluated in the first two phases. A ten-word subset of that vocabulary was used in flight to control Multi-Function Displays (MFDs) in the cockpit of an experimental F-16 jet aircraft. The MFDs contained programmable switches, which selected pages of status information or control functions. The vocabulary words enabled the pilot to either address a particular page and then a particular function on that page, a specific function on a specific page, or select an aircraft master mode. These functions could be selected either manually or by voice. Performance was approximately 90% initially, but increased to the high 90's, for some pilots, by the end of flight tests. For those pilots with performance in the high 90's, speech was the preferred mode for interacting with the MFDs. Those pilots with performance in the low 90's preferred the manual mode of operation [1].

A Tornado GR1 has been used by the UK Defence Research Agency in two series of trials, in 1989 and 1993. The 1989 trials were aimed solely at collecting speech recordings in a cockpit environment representative of modern fast jets, but a recognizer was fitted to the aircraft to provide recognition feedback to the subject. These recordings were subsequently used to assess and optimize the Marconi ASR1000 flightworthy speech recognizer.

The second series of trials was intended to demonstrate the performance of the recognizer under realistic flight conditions. The navigator's main interface to the aircraft's main computer is via the Television Tabular display, known as TV-TABS for short. This has a small keyboard, but uses a complicated menu structure to access about 40 functions. Even quite simple operations may require many key presses, and the system is difficult to use and unpopular with the aircrew. A simple physical interface to the aircraft was possible, by breaking into the keyboard bus and making the recognizer output mimic key presses. This also allowed manual input to be mixed with voice input, even within the same command. Unfortunately, software reliability problems were encountered which could not be solved in time for the flight trials. Nevertheless, a total of 19 flights were made, with the navigator reading lists of command phrases and digit strings. An average recognition accuracy of over 95% was achieved. The final vocabulary size was 99 words and the syntax had a mean branching factor of about 15.

The U. S. Air Force has been conducting in-flight tests in recent years in a NASA OV-10 aircraft. These tests are to determine the present performance of speech recognition systems in the cockpit environment. The generic task selected was controlling communications and navigation functions. The vocabulary consists of 53 words or phrases. The system was tested in flight conditions of 1g and 3g and noise levels from 95 to 115 dB. Performance levels of better than 97% were obtained for 12 subjects in these conditions, using a commercially available speaker-dependent continuous speech recognition system.

The French Delegation Generale pour l'Armement (DGA) has been supporting studies and experiments dedicated to speech recognition since 1983. From 1983 till 1989, in-flight tests (mainly on Mirage IIIB but also on Rafale-A) have pointed out speech recognition systems limitations when used in a military aircraft cockpit. In light of these results, new algorithms have been developed and experiments in a centrifuge have been conducted in order to reduce the effects under adverse conditions (noise and G-load effects: see paragraph 3.1 and 3.2). In 1989, a database was recorded during real flights under G-load on a Mirage IIIB aircraft. This database was used to evaluate speech processing and recognition algorithms performance (see [2] and [3]) before tests during real flights on the AlphaJet (described later in this section). The vocabulary was a restricted one, involving 36 words, allowing 9 linked words. The speech recognition algorithm was the preliminary version of TopVoice, the Sextant Avionique Speech

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

Table I Effects of G on speech recognition performance

Speaker - experimental conditions	Sentence Recognition Rates
Speaker 1a, 2g	100%
Speaker 1a, 4g	95%
Speaker 1b, 5g	96.4%
Speaker 1b, 2g	90%
Speaker 2, 2g	91.6%
Speaker 2, 4g	76.3%

Recognition system (previously named DIVA). This speech recognition system is speaker-dependent, based on Dynamic Time Warping pattern recognition.

Two speakers took part in these experiments. Speaker 1 appears twice (Speaker 1a and Speaker 1b) because he used two different oxygen masks. The results are shown in Table I.

Remark : For the results described below (real flights in an AlphaJet), the recognition rate is a Sentence Recognition one: a whole sentence is considered as misrecognized as soon as there is only one recognition error, whatever the error is (deletion, substitution or insertion).

After these preliminary database experiments, TopVoice has been tested during flights on AlphaJet, the French training jet.

All flight configurations have been tested (speed from 200 to 450 knots, flight phases under G-load effects, low flight levels,

Table II Sentence recognition rate, including all flights all speakers

First utterance	90%
First repetition (in case of error on the first utterance)	95%
Third utterance (in case of error on the first repetition)	97%

real commands in context). Two syntaxes were defined in order to take into account new functionalities involved in modern military fast jets (example: Rafale). The first one, to be used during cruise flight phases, involved more than 150 words and allowed sentences whose maximum length was about 10 words. The second one was designed for flight under G-load and contained 25 real-time commands. These evaluations consisted of 80 flights, involving 15 different speakers and more than 10,000 vocal commands to recognize. The results (see Tables II and III) show that the Sentence Recognition Rate (SRR) increases as soon as the pilot's attention increases.

These evaluations are broad enough to draw conclusions about the main parameters that influenced the performance:

- Noise is obviously one of these parameters, since the sentence recognition rate decreases as the noise level increases. Note that noise level increases not only during flight phases under G-load, but also as the speed increases. Despite this noise level, noisy speech processing avoids bad recognition rates and appears efficient.
- The microphone and audio circuitry must be optimized.
- The different parts of the syntax do not lead to the same results: systems commands, isolated words, and digits are well recognized, but international alphabet or numbers appear to be more difficult to recognize.
- Speech recognition remains very tied to speaker habituation as the results show. It depends on training phase quality and speaker vocal characteristics.

On the other hand, some effects are not so relevant as it seemed; specifically, G-load and Lombard effects.

The subjective conclusions of the users were that it appears easier to obtain data and parameters from the system when using vocal command. With a more and more complicated system to manage, vocal command is a relevant tool to decrease the workload, but vocal command must be controlled by a system able to detect recognition errors and to avoid disastrous consequences of speech recognition mistakes. Does it induce a dialogue between the pilot and the system, as soon as an error is detected? And first of all, how to detect erroneous recognitions?

Such evaluations have shown the technical feasibility of speech recognition during flight and have identified some operational problems to solve, the main one being the system's ability to control its own recognition and to manage erroneous recognition.

The U. S. Air Force and the U. S. Navy are also conducting flight tests of speech recognition in the Joint Strike Fighter program. The application is a means of managing information and sensors. The vocabulary for this application is 12 words. The system has been tested in operational flight test conditions. Performance levels of 70% or greater were obtained with three pilots. Two of the three pilots had performance of 90% or greater on several flights. The system tested was a militarized speaker-dependent isolated-word speech recognition system.

The European Fighter Aircraft EF2000 is a single-seat agile combat aircraft, planned to enter service about 2002. Speech input was included in the requirement from the beginning, and will be used for control of displays, radar, radios, target designation, navigation aids, and several other functions. Although test flying of the aircraft commenced in 1994, development of the speech recognizer module has not reached the stage of flight trials (at the time of writing). A commercially available speech recognizer has, however, been integrated into the cockpit simulator and used in the development of the man-machine interface. The reaction of pilots during the assessment program has been very positive. They regard speech recognition as essential to the safe and

Table III Sentence Recognition Rate under G-load effects (5g)

First utterance	90%
First repetition (in case of error on the first utterance)	94%
Third utterance (in case of error on the first repetition)	98%

efficient operation of the aircraft.

2.2. ROTARY WING

The first in-flight use of ASR in a helicopter was in January 1981 [4]. These tests demonstrated that the most important problem to overcome for ASR in helicopter applications is the high noise level during flight.

The Day/Night All Weather (D/NAW) program in the UK, and the associated Covert Night and Day Operations in Rotorcraft (CONDOR) collaboration between the USA and the UK, are primarily concerned with advanced visual systems to allow rotary-wing operations to proceed in very poor visibility. The reliance on helmet-mounted displays can create a problem for the aircrew in operating switches and controls inside the aircraft, so voice input is an important adjunct to the visually coupled system.

Preliminary recognition trials in the DERA noise and vibration simulator have given good results. Mission-based trials in the Helicopter Mission Simulator in January 1997 compared missions flown with and without the use of voice input. Both pilot and commander had voice input, with different vocabularies. The pilot used about 25 words to control display modes and the radio altimeter (radalt); the commander's vocabulary of about 45 words controlled radios, map displays, transponder, and radalt. After the trial, the subjects, mainly operational Army aircrew with no previous experience of voice input, were strongly in favor of it, and considered it would offer a considerable enhancement to mission effectiveness. Following the simulator trials, a commercial speech recognizer was installed on the Lynx helicopter used for the D/NAW program at DERA, Boscombe Down, in the UK. Flight tests in late 1997 gave over 98% word accuracy.

Speech recognition has been tested on Gazelle, as a component of a Real-Time Digital Map Generator named MultiHélicare provided with graphical symbology overlaying capabilities. MultiHélicare is connected to the aircraft navigation system, to a voice command system and to a transmission system. The operator controls MultiHélicare with a joystick and the voice command system. The voice command system is TopVoice provided by Sextant Avionique, and which was described in the previous section.

Actions of the operator allow management of:

- the underlying map presentation,
- the overlaying symbology presentation,
- the loading and saving of the mission data,
- the aircraft navigation,
- the communications with another MultiHélicare system.

The main functionalities of MultiHélicare are:

- the friends/enemies tactical situation presentation and modification,
- the flight plans presentation and modification with automatic guidance,
- the dynamic terrain analysis by coloration and profiles display.

The syntax used for this application involved 67 French words and 2150 possible different sentences. The average length of the sentences is 3.5 words and the branching factor is 6.3. The SRR is over 95% during real flights, for any pilot. Moreover, tests conducted using an equivalent German syntax resulted in an SRR of over 98%.

The system has been used for several months during real flights, and it is very important to point out the subjective appreciation of the users who consider that the integration of speech recognition in a system such as MultiHélicare provides a tremendous amount of increased abilities, while decreasing the workload.

2.3. SPACE

Investigations into the utility of voice input/output (I/O) in the space shuttle were initially conducted in the mid 1980's [5]. The investigations centered on the control of the shuttle's Multifunction Cathode ray tube Display System (MCDS). This system is the main method the astronauts have for interacting with the five flight computers. Through the MCDS system, the astronauts do everything from reconfiguring the flight computers to checking the mission elapsed time. The MCDS has a 32-key oversized keyboard designed for use with the bulky gloves of a space suit. A commercially available, speaker-dependent speech recognition system was used as an alternative to the keyboard. Similar applications of voice I/O are being considered for the space station as well [6].

An experimental voice command system was carried on shuttle mission STS-41 in October 1990, with the aim of collecting data on speech in microgravity conditions and to demonstrate the operational effectiveness of controlling spacecraft systems by voice. The recognizer was interfaced to the orbiter's closed circuit TV system, which allows the astronauts to monitor the payload bay from inside the flightdeck. The speaker-dependent system used a vocabulary of 41 words to control the four TV cameras mounted in the payload bay. A very simple syntax allowed the cameras to be panned, tilted, focused, and allocated to one of two monitors. Two astronauts used the speaker-dependent system, with templates created on the ground before the mission. The system had the capability to retrain templates in space should the need arise. One of the astronauts experienced some initial difficulties due to the placement of his microphone, which was boom-mounted on a very lightweight headset. Once this was corrected, the system gave good results, and both astronauts were pleased [7].

There are plans for further assessment of voice input on future shuttle flights, possibly using it to control the manipulator arm. As a preliminary, the Canadian Space Agency included an experiment on simulated voice control of a robot arm during a short-duration space mission simulation. Four trainee astronauts spent seven days isolated in a hyperbaric chamber with workload and living conditions similar to those encountered in space, except for the gravity. The voice control tasks consisted of instructing a simulated 6 degree-of-freedom manipulator arm to grasp a ball while avoiding obstacles. The voice recognizer was simulated with the "Wizard of Oz" technique. The astronauts were not given a fixed vocabulary other than starting each command with the word "Viktor," but spoke spontaneously. Despite this, they used only 107 words in total between them, and only about 30 of these were common to all speakers. This experiment has helped to identify the vocabulary and syntax most natural for the task and will contribute to further evaluation of voice input in space applications.

2.4. COMMAND AND CONTROL

In the late 1980's, researchers at Boeing [8] investigated the utility of speech input/output (I/O) in the Airborne Warning and Control System (AWACS) man/machine interface. The present AWACS interface provides control and management of sensors through updating fields in tabular displays by inserting or changing alphanumeric values. This interface proves adequate for controlling one or two sensors, but it begins to overload the operator as more sensors are added. Operator tasks were analyzed to identify those thought to be best performed by speech I/O. Based on these functions a vocabulary and grammar were developed for a commercially available speech recognition system. This system demonstrated the effectiveness of voice I/O for several functions including fuel updating, committing fighters, and tactical broadcast control. The studies identified several features required of speech I/O in the AWACS operational environment. In the AWACS environment, the operator is under stress, there are multiple voice communications occurring in the background, and few I/O errors (speech input not recognized, speech output not heard by the operator) can be tolerated.

Another application of speech recognition in the late 1980's was training of air traffic control (ATC) trainees in the use of the correct ATC technology and phraseology [9]. The concept is that the trainee/speaker runs through a set of ATC scenarios. He speaks sentences intended to be appropriate to the scenario. The system provides feedback, identifying items and places where the vocal behavior of the trainee must be altered. The trainer used a commercially available, speaker-dependent continuous speech recognition system.

2.5. MONITORING

It is common experience that many aspects of a person's physical or emotional state may be detected from the sound of his voice, but detailed knowledge relating changes in measurable parameters of speech to particular kinds of stress is very limited. Two major problems are that stress can be very difficult to define, and that individual reactions to it may vary over a very wide range. However, given that humans can classify others' emotional states from their voices with some degree of accuracy, it must be possible, at least in principle, to automate the process.

Physical stresses, such as G-force and vibration, have relatively well defined effects on speech production, because they act directly on the vocal apparatus without the mental interpretation that intervenes in the case of many other stressful stimuli. Nevertheless, the effects are still dependent on the subject's level of training and experience under the particular stressor. In practice, the physical conditions in an aircraft can be measured accurately and reliably by physical sensors, so there is little need to use voice monitoring in this way. It may, however, find an application in accident investigations when there are no physical measures available.

There is considerable psychological literature on the effects of stress and emotion on the voice [10], but most of the practical interest has been associated with space flight. Given the isolation, danger and expense of space missions, monitoring of the astronauts' state may be crucial to avoiding a disaster. Stress levels can be determined by means of physiological measures, but the associated sensors and wiring will be inconvenient in the confined cabin of a spacecraft. Also, where the astronaut is required to work outside the spacecraft, they will complicate the process of donning the pressure suit, and will require extra telemetry bandwidth. There has therefore been considerable interest in using the voice to monitor the state of the astronaut.

Many experiments have found changes occurring to voice parameters under stress conditions, but there have always been very large differences in responses between speakers. Some of this variance is associated with different reactions obtained from different personality types, and also with gender. It is possible, though, that more consistent reaction would be obtained from a group as highly selected and trained as astronauts are. Even so, a reliable "voice stress monitor" seems a long way off.

3. APPLICATION PROBLEMS

Speech processing's influence on speech recognition performance is obvious. Classical speech processing can be improved by various algorithms (speech/noise discrimination, denoising algorithms,...) whose aim is to take into account the particular environmental characteristics of military applications. This next section will address several such techniques. Subsequent sections will discuss other challenges facing the application of speech-based control in aerospace environements.

3.1. NOISE

One problem that all current systems share is that their performance degrades significantly as conditions depart from the ideal noise-free case. Recognition errors can increase dramatically in the presence of noise, which can come in a variety of forms. All real-world applications are subject to interference from noise, whether it is due to fans in an office environment, vehicle engines, machinery, or even other voices in the background. The usefulness of an ASR system is limited by how well it can handle such problems.

Recently, there has been some success in this area. Two areas of focus have emerged: the use of additional knowledge sources such as improved language models or prosody and improved modeling of the acoustic phenomena.

Woods asserts that "there is not enough information in the acoustic signal alone to determine the phonetic content of the message". Humans rely on other knowledge sources to help constrain the set of possible interpretations. These are useful for machine recognition as well, and indispensable for many tasks. The most important knowledge source is grammar. Grammar places strong restrictions on the set of words which can follow or precede a given word [11-15]. Others include prosody (information contained in the rhythms and pitch variations of speech) [16, 17] and focus (constraining the vocabulary to the topic of a "conversation") [12].

Modeling of acoustic phenomena has focused primarily on reducing the effects of noise on speech. To combat this kind of effect, various speech enhancement techniques have been investigated. These have resulted in error reduction in ASR systems of around 35% to nearly 100%, depending on the task and the amount of noise [18-24].

Another area of focus has been that of representing the acoustic signal in ways that relate to the human auditory system, since humans perform very well at speech recognition [25-28]. Since a very large reduction in data dimension and data rate takes place between the sampling of an acoustic signal and the representation of that signal which is used in the recognition algorithm, it is critical that the reduction take place in a manner

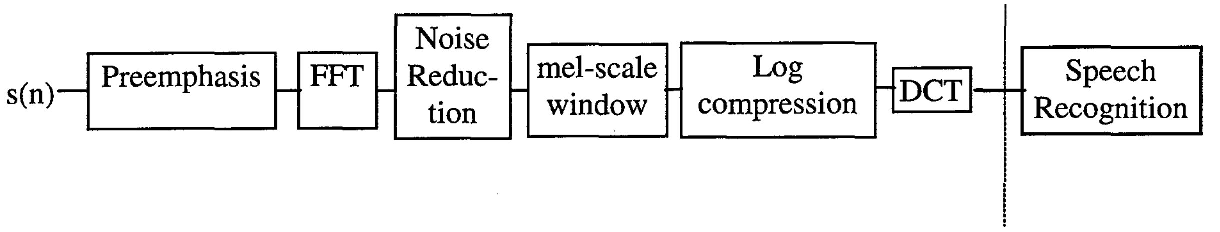


Figure 1 Speech Processing

which preserves the important information. Filters and signal processing methods are designed which mimic processes that occur in the inner ear and the brain. It is thought that the information extracted by these techniques is likely to be linguistically relevant and more robust to effects of noise.

Among these methods, some have been implemented by speech/noise discrimination [2, 3]. Noise robustness associated with speech/noise discrimination has been involved in a flyable speech recognizer which has been tested during flights on helicopters and fast jets (see sections 2.1 and 2.2). As described in [2], preliminary experiments on a database recorded during flights of a Mirage IIIB (see section 2.1.) have shown that speech detection alone was able to improve the speech recognition rate, even under G-load effects. If noise cancellation is added, there is an additional speech recognition rate gain, but which is lower than the gain due to detection.

It is quite obvious that speech/noise discrimination improves the speech recognition rate. It is more difficult to understand why speech detection is so important. In fact, a pilot uses a push-to-talk (PTT) in order to give a voice command to the system. The pilot's PTT is not perfect and, in most cases, is longer than the real speech duration. Speech recognition algorithms begin the recognition process during a noisy pause; this can induce bad choices in the syntactic tree structure. Owing to accurate speech detection through speech/noise discrimination, such a phenomenon can be avoided.

Speech/noise discrimination and noise cancellation are closely related problems, because noise cancellation algorithms need statistical and spectral information about the background noise of interest. The noise can be considered stationary during a vocal command, but from one vocal command to another, its characteristics (for example, its level) can change. So, noise cancellation requires the detection of noise to adaptively extract its spectral and statistical parameters. The ability to discriminate speech from noise enables the calibration of noise cancellation algorithms. The result of such an approach is described by Figure 1 that depicts the whole processing chain. Noise cancellation is assumed to be performed by Wiener Filtering.

This principle has been tested on a database recorded during real flights under G-load on Mirage III B (see section 2.1). The results obtained are described in Table IV, where the nomenclature is the following one:

- PTT: results obtained when the pilot's original Push-To-Talk is used in order to define the beginning and the end of the utterance
- SD: results provided with Speech Detection alone
- SD+NC: results provided by the complete algorithm (Speech Detection and Noise Cancellation)
- PWB+NC: results obtained with a Perfect Word Boundary Detection and Noise Cancellation

In each column of Table IV, the number of errors and the number of utterances are given, as well as the recognition rate: for example, 12/30 (60%) indicates 12 errors in 30 utterances, and the recognition rate is then 60%.

Figure 2 illustrates the Noise Cancellation efficiency of such an approach on the utterance "Donne Page Hydraulique."

3.2. STRESS

Stress is a rather ill defined concept, covering a multitude of generally threatening conditions. Many of these have elements in common, particularly those that activate the autonomic nervous system, but the external stimulus is always subject to a greater or lesser degree of mental interpretation which results in individual reactions varying widely. In addition, training and experience can have a large effect on how well individuals cope with many kinds of stressors. The effects of stress are

Table IV speech recognition rates with/without speech/noise discrimination and with/without noise cancellation

Environmental conditions	PTT	SD	SD+NC	PWB+NC
Speaker 1 - 2g	5/36 (86.1%)	2/36 (94.4%)	0/36 (100%)	0/36 (100%)
Speaker 1 - 4g	4/60 (93.3%)	5/60 (91.6%)	3/60 (95%)	3/60 (95%)
Speaker 1 - 5g	3/28 (89.2%)	4/28 (95.1%)	1/28 (96.4%)	1/28 (96.4%)
Speaker 1 - 2g	12/30 (60%)	6/30 (80%)	3/30 (90%)	2/30 (93.3%)
Speaker 2 - 2g	39/48 (18.75%)	11/48 (77%)	4/48 (91.6%)	2/48 (95.8%)
Speaker 2 - 4g	53/55 (4%)	23/55 (58%)	13/55 (76.3%)	8/55 (85.4%)

usually apparent in the voice, and hence affect the performance of speech recognizers. The problem, as always, is that the conditions of use are different from those under which the recognizer's models are trained. This mismatch is largely unavoidable, as it is usually impractical, expensive or unethical to subject a user to such stresses in order to train the recognizer.

3.2.1. PHYSICAL STRESS

Physical stresses may be classified under four main areas: the force environment, auditory distraction, the thermal environment, and personal equipment. For aircrew, the major factors in the force environment are G-force, vibration and pressure (cabin pressure or pressure breathing for G protection). Some experiments have shown that highly trained and experienced personnel can speak relatively normally at up to 5g with only about 5% loss in recognizer performance, but inexperienced subjects may suffer 30% loss in performance at lower G-levels. Vibration is the predominant problem in rotary wing aircraft. Dominant frequencies from the main rotor lie in the range of 5-30 Hz; typical resonant frequencies of body structures of the torso and head also lie in this range. Pressure breathing for G-protection involves increasing the pressure of the breathing gas by as much as 50 mmHg or more. This inflates the vocal tract and makes speaking difficult.

Some studies have been conducted in order to determine not only the speech recognition rate degradation due to G-load effects, but also in order to point out efficient speech processing able to balance these degradations owing to an analysis of speech production alterations under G-load.

These studies are based on experiments in a centrifuge, involving six pilots whose mean age was 30. Through different signal analysis tools (pitch detection, short time Fourier transform, Multiresolution analysis, Principal Component

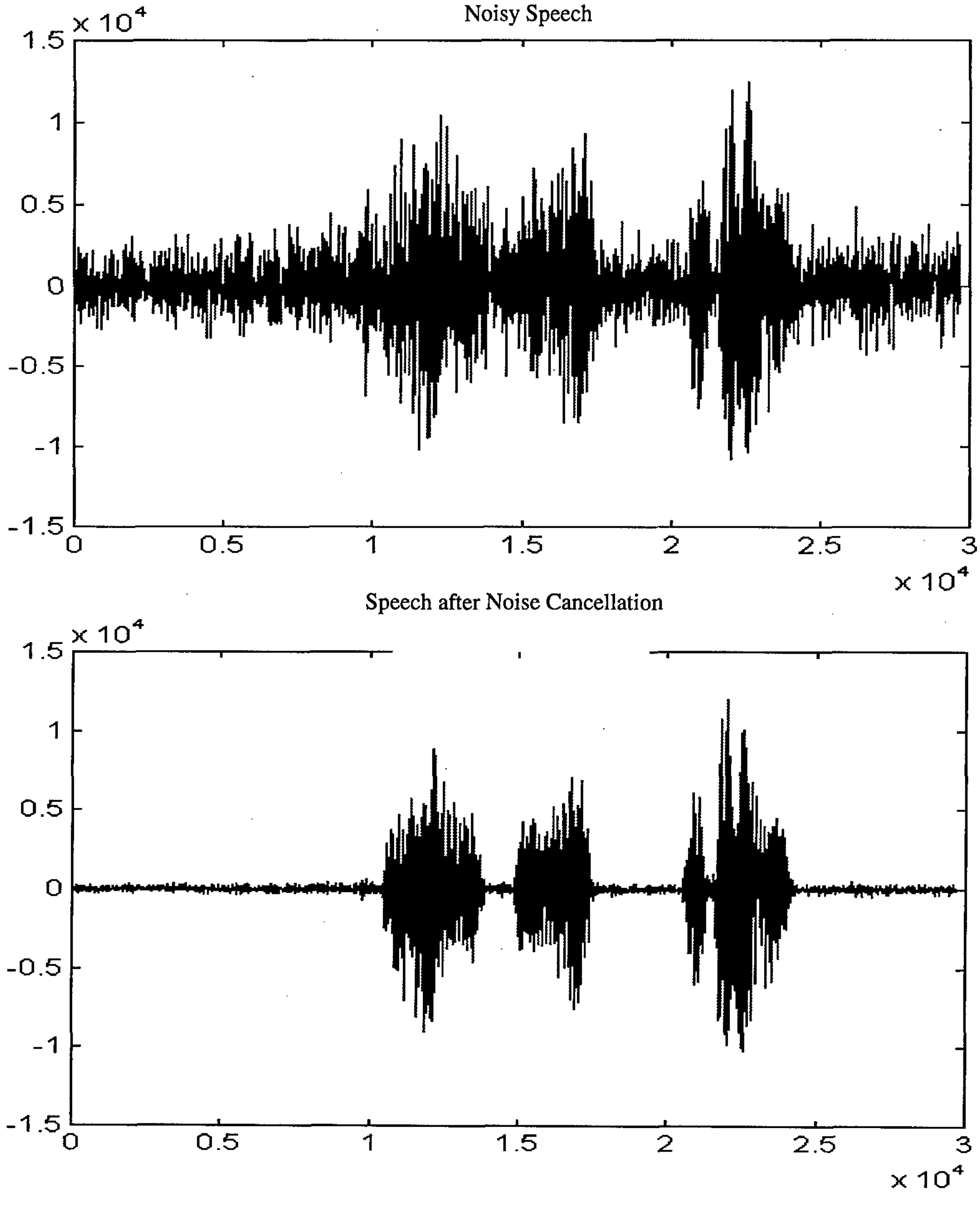


Figure 2 Noisy speech (top) and after Noise Cancellation (bottom)

Analysis), it has been possible to study speech production modifications at different G-load levels (1.4g, 3g, 6g). Even if these tools point out some typical phenomena correlated with identified physiological mechanisms, it remains difficult to integrate such considerations in a speech recognition system since these phenomena remain variable and hazardous.

However, this study points out that detecting speech from pause and reducing the vocabulary complexity were two relevant means in order to get acceptable speech recognition, even under G-load effects. Speech detection principles and influences on the speech recognition task are, under G-load effects, the same as those described in section 3.1. Reducing syntax complexity is not simply a trick but fits with section 4 recommendations and human physiological abilities, since it becomes really difficult to speak clearly and naturally, except for highly trained personnel. Finally, from an operational point of view, the number of required speech commands decreases quickly.

In order to take into account each environmental parameter whose influence is relevant, some studies have been conducted in order to determine the speech production modifications due to combined stress (workload, noise, G-load, positive pressure breathing). A database has been recorded in a centrifuge and the data are currently being processed. Such an analysis should provide some constraints that future speech recognizers will have to respect in order to be relevant under complex fast jets environmental conditions. Such a theme is close to the current NATO working group IST/TG001 (formerly RSG10) dedicated to state-of-the-art speech processing.

Noise levels are high in modern military aircraft, often 110-115 dB SPL. Hearing protection is improving, but many aircrew can still expect to be subject to levels of around 85 dBA for the duration of the mission. Short-term effects can be compensated by training the recognizer under similar noise conditions, but these noise levels can also create mental fatigue over a period. Other auditory stressors include auditory warnings and voice communications that add to the total noise dose and may carry distracting or anxiety-causing information.

The thermal (i.e. temperature and humidity), environment of military aircraft is in general not too extreme, but may become so in the event of a failure or battle damage. At present, there is not much detailed knowledge about the effects of temperature on the voice.

Personal equipment includes clothing, helmet, oxygen mask, NBC protection, and safety harnesses. These may restrict movement in various ways or apply pressure to the body. The oxygen mask is a special case, in that it is intimately involved in speech production. The effect that the mask has on the speech spectrum is considerable, but is not a stressor as such. The mask may also constrict jaw movement, add to fatigue, and, over a long period, apply painful pressure to the face.

3.2.2. EMOTIONAL STRESS

Emotional stresses may be classified under the general headings of task load, mental fatigue, mission anxieties and background anxieties. Task load arises out of the immediate demands of the mission on a crewmember, requiring him to absorb information, make decisions and take actions. Mental fatigue affects general alertness, and may arise from loss of sleep, physical fatigue or boredom. Mission anxiety arises out of threatening situations that occur in the course of the mission. As well as the obvious threats arising from enemy action, this also covers social aspects such as the weight of responsibility and difficulties in interactions between crewmembers. Finally, background anxieties covers aspects of domestic, career and health worries that do not arise out of the mission itself but can have a significant impact on aircrew performance.

3.3. ACCENT

It is well known that speaker accent is one factor that degrades the performance of present-day speech recognition systems [29]. This is a problem that occurs no matter the target language on which the recognizer was trained [30]. Approaches to this problem are to first identify the accent [31-33] and then use a recognizer trained on that accent [34, 35], select an appropriate language model [36], or adapt to the accent/speaker [37]. Each of these approaches has trade-offs in terms of training complexity.

Degradation in recognition performance due to accent is a concern in commercial applications running on the telephone network and on personal computers. It is also a concern in military applications with the now-common multinational forces and in air traffic control. This area will get increased attention because of the significant benefits that will be derived in commercial applications. The unique military aspects will be the effects on speech recognition performance with combinations such as accented speech in a stressful, high-noise environment.

4. REQUIRED ENHANCEMENTS

Speech recognition performance for small and large vocabulary systems is adequate for some applications in benign environments. Any change in the environment between the training and testing causes degradation in performance. Continued research is required to improve robustness to new speakers, new dialects, and channel or microphone characteristics. Systems that have some ability to adapt to such changes have been developed [38, 39]. Algorithms that enable ASR systems to be more robust in noisy changing environments such as airports or automobiles have been developed [40-43], but performance is still lacking. Speech recognition performance for very large vocabularies and large perplexities is not adequate for applications in any environment. Continued research to improve out-of-vocabulary word rejection in addition to the above-mentioned areas will enable larger vocabulary ASR systems to be viable for applications in the future.

An answer to the problem of the user having to remember a large vocabulary is to make the system capable of understanding any command, however it is phrased. The user can then speak naturally, using whatever form of words comes to mind at that instant. This removes the workload associated with having to remember which words are valid. Such systems are often called "speech understanding" systems.

The simplest systems use word-spotting techniques. For example, to select a radio frequency with a finite state syntax, the pilot may have to say, "RADIO VHF HEATHROW APPROACH." A natural language system could accept "GIVE ME HEATHROW APPROACH ON VHF" or "SELECT VHF, ER, I WANT HEATHROW APPROACH." The system needs only recognise the words "VHF," "HEATHROW," and "APPROACH" to infer that the VHF radio should be tuned to that channel. Words which are not a good match to keywords in the vocabulary are matched to a so-called "garbage model,"

which approximates the long-term speech spectrum. Another approach is to attempt to recognize all words spoken, then pick out the key words from the resulting word stream. The overall error rate may be relatively poor, but providing that the key words are recognized correctly, useful output may be obtained.

Many speech understanding systems attempt to make use of several different areas of knowledge about the speech and the situation in which it is being used. Starting with a parametric representation of the speech signal, hypotheses are formed about possible phone sequences. Phonetic and phonological knowledge is used to provide constraints at this level. From these sequences, higher-level hypotheses are formed about possible word sequences using syntactic, prosodic and lexical knowledge. Constraints may be added from knowledge of the application and the current situation, until finally a single sentence emerges. A reliable natural language interface may be some way off, but is a prime goal for research in speech recognition.

Use of speech-based control as a supplement to conventional controls is becoming common. For example, a system designer can make cockpit radio frequency selection or multi-function display operation accessible with speech-based as well as conventional control systems. The user could choose to use the speech-based system when appropriate. An analogy is the availability of both keyboard and mouse functions for cursor positioning in a modern personal computer system. Users will choose one or the other depending on the nature of the task, hand location and personal preference. One key issue that must be addressed is the ability to operate speech-based controls in multi-task environments. Some research has investigated the effect of task loading and other physical stressors on speech and its resultant impact on speech recognition performance [44, 45]. Continued research is needed to reduce the impact of these factors.

5. REFERENCES

1. Howard, D., "Flight testing of the AFTI/F-16 voice interactive avionics system", in "Proc. of Military Speech Technology", 1987, pp 76-82.

2. Pastor, D. and Gulli, C., "Improving Recognition Rate in Adverse Conditions by Detection and Noise Suppression", in "Proc. of ESCA Workshop on Speech Recognition in Adverse Conditions", 1992.

3. Pastor, D. and Gulli, C., "DIVA 5 Dialogue Vocal pour Aeronef: Performance in Simulated Aircraft Cockpit Environments", in "Proc. of Joint ESCA-NATO/RSG10 Tutorial and Workshop: Application sof Speech Technology", 1993.

4. Simpson, C.A., Coler, C.R., and Huff, E.M., "Human Factors of Voice I/O for Aircraft Cockpit Controls and Displays", in "Proc. of Workshop on Standardization for Speech I/O Technology", 1992.

5. Hoskins, J.W., "Voice I/O in the Space Shuttle", in "Speech Technology", 1984, pp 13-18.

6. Castigione, D. and Goldman, J., "Speech and the Space Station", in "Speech Technology", 1984, pp 19-27.

7. Salazar, G., "Voice Recognition Makes a Debut on NASA STS-41 Mission", in "Speech Technology", 1991.

8. Salisbury, M. and Chilcote, J., "Investigating Voice I/O for the Airborne Warning and Control System (AWACS)", Speech Technology, 5, 1, 1989, pp 50-55.

9. Benson, P. and Vensko, G., "A Spoken Language Understanding System for Phraseology Training of Air Traffic Controllers", Speech Technology, 5, 1, 1989, pp 64-69.

10. Scherer, K.R., "Voice, Stress, and Emotion", in "Dynamics of Stress", M. Appley and R. Trumbull, Editors: New York, 1986, pp 157-179.

11. Woods, W.A., "Language processing for speech understanding", in "Readings in Speech Recognition", A. Waibel and K.-F. Lee, Editors, Morgan Kaufmann: San Mateo, CA, 1990, pp 519-533.

12. Young, S.R., Hauptman, A.C., Ward, W.H., Smith, E.T., and Werner, P., "High-level knowledge sources in usable speech recognition systems", in "Readings in Speech Recognition", A. Waibel and K.-F. Lee, Editors, Morgan Kaufmann: San Mateo, CA, 1990, pp 538-549.

13. Lee, K.-F., "Large-Vocabulary Speaker-Independent Continuous Speech Recognition: The SPHINX System", Ph.D. Thesis, Carnigie Mellon University, 1988.

14. Lee, K.-F., Hon, H.-W., and Reddy, R., "An overview of SPHINX speech recognition system", IEEE Trans. Acoust., Speech, and Signal Processing, 38, 1, 1990, pp 35-45.

15. Lee, K.-F., "Context-dependent phonetic hidden Markov models for speaker-independent continuous speech recognition", IEEE Trans. Acoust., Speech, and Signal Processing, 38, 4, 1990, pp 599-609.

16. Waibel, A., "Prosody and Speech Recognition", Ph.D. Thesis, Carnegie Mellon University, 1986.

17. Waibel, A., "Prosodic knowledge sources for word hypothesization in a continuous speech recognition system", in "Readings in Speech Recognition", A. Waibel and K.-F. Lee, Editors, Morgan Kaufmann: San Mateo, CA, 1990, pp 534-537.

18. Ephraim, Y., Malah, D., and Juang, B.-H., "On the application of hidden Markov models for enhancing noisy speech", IEEE Trans. Acoust., Speech, Signal Processing, 37, 12, 1989, pp 1846-1856.

19. Erell, A. and Weintraub, M., "Estimation using log-spectral distance criteria for noise robust speech recognition", in "Proc. of Intl. Conf. Acoust., Speech, and Signal Processing", 1990, pp 853-856.

20. Varga, A.P. and Moore, R.K., "Hidden Markov model decomposition of speech and noise", in "Proc. of Intl. Conf. Acoust., Speech, and Signal Processing", 1990, pp 845-848.

21. Acero, A. and Stern, R.M., "Environmental robustness in automatic speech recognition", in "Proc. of Intl. Conf. Acoust., Speech, and Signal Processing", 1990, pp 849-852.

22. Ephraim, Y., "A Bayesian estimation approach for speech enhancement using hidden Markov models", IEEE Trans. Signal Processing, 40, 4, 1992, pp 725-735.

23. Boll, S.F., "Speech enhancement in the 1980s: Noise suppression with pattern matching", in "Advances in Speech Signal Processing", S. Furui and M. Sondhi, Editors, M. Dekker: New York, 1992, pp 309-325.

24. Ephraim, Y., "Gain-adapted hidden Markov models for speech recognition of clean and noisy speech", IEEE Trans. Signal Processing, 40, 6, 1992, pp 1313-1316.

25. Williams, C.S., "Designing Digital Filters", Englewood Cliffs: Prentice-Hall, 1986.

26. Tishby, A., "A dynamical systems approach to speech processing", in "Proc. of Intl. Conf. Acoust,, Speech and Signal Processing", 1990, pp 365-368.

27. Ghitza, O., "Auditory nerve representation as a basis for speech processing", in "Advances in Speech and Signal Processing", S. Furui and M. Sondhi, Editors, M. Dekker: New York, 1992, pp 453-485.

28. Junqua, J.-C., Wakita, H., and Hermansky, H., "Evaluation and optimization of perceptually based ASR front end", IEEE Trans. Speech and Audio Process., 1, 1, 1993, pp 39-48.

29. Pallett, D. and Fiscus, J., "1996 Preliminary Broadcast News Tests", in "Proc. of DARPA Speech Recognition Workshop", 1997.

30. Kudo, I., Nakama, T., Watanabe, T., and Kameyama, R., "Data collection of Japanese dialects and its influence into speech recognition", in "Proc. of Intl. Conf. on Spoken Language Systems", 1996, pp 308-311.

31. Hansen, J.H.L. and Arslan, L.M., "Foreign Accent Classification using source generator based prosodic features", in "Proc. of Intl. Conf. on Acoust., Speech, and Signal Processing", 1995, pp 836-839.

32. Kumpf, K. and King, R.W., "Automatic accent classification of foreign accent Australian English speech", in "Proc. of Intl. Conf. on Spoken Language Systems", 1996, pp 4-7.

33. Teixeira, C., Trancoso, I., and Serralheiro, A., "Accent Classification", in "Proc. of Intl. Conf. on Spoken Language Systems", 1996, pp 577-580.

34. Arslan, L.M. and Hansen, J.H.L., "Language accent classification in American English", Speech Communication, 18, 4, 1996, pp 353-367.

35. Arslan, L.M. and Hansen, J.H.L., "Improved HMM triaining and scoring strategies with application to accent classification", in "Proc. of Intl. Conf. on Acoust., Speech, and Signal Processing", 1996, pp 598-601.

36. Humphries, J.J., Woodland, P.C., and Pearce, D., "Using accent-specific pronunciation modelling for robust speech recognition", in "Proc. of Intl. Conf. on Spoken Language Systems", 1996, pp 623-626.

37. Diaoloukas, V., Digalakis, V., Neumeyer, L., and Kaya, J., "Development of dialect-specific speech recognizers using adaptation methods", in "Proc. of Intl. Conf. on Acoust., Speech and Signal Processing", 1997, pp 1455-1458.

38. Lee, C., Lin, C.-H., and Juang, B.-H., "A study on speaker adaptation of the parameters of continuous density hidden Markov models", IEEE Trans., 39, 4, 1994, pp 806-814.

39. Neumeyer, L. and Wientraub, M., "Robust speech recognition in noise using adaption and mapping", in "Proc. of Int. Conf. on Acoust., Speech and Signal Processing", 1995, pp 141-144.

40. Acero, A. and Stern, R.M., "Robust speech recognition by normalization of the acoustic space", in "Proc. of Intl. Conf. on Acoust., Speech and Signal Processing", 1991, pp 893-896.

41. Hermansky, H., personal communications, 1995.

42. Hirsch, H., Meyer, P., and Reuhl, H.W., "Improved speech recognition using high-pass filtering of subband envelopes", in "Proc. of Eurospeech", 1991, pp 413-416.

43. Murveit, H., Butzberger, J., and Weintraub, M., "Reduced channel dependence for speech recognition", in "Proc. of DARPA Speech and Natural Language Workshop", 1992, pp 280-284.

44. Stanton, B.J., "Robust recognition of loud and Lombard speech in the fighter cockpit environment", Ph.D. Thesis, Purdue University, 1988.

45. Rajasekaran, R.K. and Doddington, G., "Robust speech recognition: initial results and progress", in "Proc. of DARPA Speech Recognition Workshop", 1986, pp 73-80.

TECHNOLOGY AND APPLICATION OF HEAD BASED CONTROL

Joshua Borah
Applied Science Laboratories
175 Middlesex Turnpike
Bedford, MA 01730

1. SUMMARY

This lecture reviews the use of head position and orientation as a means for human interaction with computers and other systems, especially in the military aerospace environment. It addresses the reasons for using head based control, current measurement technology, relevant physiological and behavioral factors, and the uses of head based control to date.

2. REASONS FOR CONSIDERING HEAD BASED CONTROL

People normally direct their visual attention by facing their head toward the general area of interest, and by using eye motion to focus more finely on areas within the central field of view.

Head motion based control attempts to take advantage of this natural behavior in order to facilitate tasks that would otherwise take longer and occupy other manual and cognitive resources, or to increase the richness of information that can be presented by using knowledge of where attention is focused.

Head tracking instrumentation is a mature technology that has already seen significant operational service in military aircraft. Past and current uses include designation of external targets for weapons delivery systems, and slaving of external airframe mounted sensors, such as radar and thermal sensors. The former (target designation) requires that the target be sighted through a head mounted aiming reticule, and is an example of explicit control. The head is purposely positioned to affect a control input. The later (slaving of an external sensor) is an example of implicit control. No special head motion task is required. The pilot simply moves his head naturally, but enhanced information corresponding to the pilot's central field of view can be continually presented on a head mounted display.

The example of a head slaved sensor is a case in which a head mounted display image is made to appear stable with respect to the outside environment. Head position measurement is also required if a head mounted display image must appear to be stabile with respect to the cockpit. Future use of virtual environments, for example, may require that images of virtual controls created by a helmet mounted display appear to be fixed to the airframe.

In the future, eye line of gaze, rather than just head position, may be used to designate targets, or to interact with objects and switches in the cockpit or in virtual environments. If eye position is measured with respect to the headgear, as it probably will be, head position and orientation measurement is still required to determine line of gaze with respect to the airframe. Thus a head tracker will usually be an integral part of any line of gaze measurement system.

3. METHODS FOR MEASURING HEAD POSITION

The predominant techniques for measuring head position and orientation can be classified as mechanical, inertial, acoustic, optical, and magnetic. Mechanical, optical, and magnetic head tracking techniques have already seen operational use in military aircraft. In recent years magnetic systems have probably seen the widest use and can be considered a relatively mature technology for the aerospace environment. Although some specific implementations have been designed to measure only head orientation, all categories of system can theoretically measure all 6 degrees of freedom.

Translation measurements (3 degrees of freedom) specify the location of a fixed point on the head gear with respect to a fixed origin in the airframe. Translation is typically specified in Cartesian coordinates, but can also be specified in polar coordinates. Orientation measurements (3 degrees of freedom) specify the orientation of a coordinate frame that is fixed to the head gear relative to coordinates that are fixed to the airframe. Orientation is typically specified as 3 Euler angles, a 9 element rotation matrix, or a set of 4 quarternions.

Head tracker performance is often described in terms of some of the following parameters, usually specified separately for translation and orientation measures. *Accuracy* is the expected difference between measured position and true position. *Precision* (repeatability) is the expected difference in repeated measurements of the same true position. *Resolution* is the smallest change in true position that can be reported by the device. *Range* is the maximum excursion from some specified nominal position over which valid measurements can be made. Orientation range is usually specified in terms of the three Euler angles, and translation range is usually specified as a three dimensional region of space ("motion box"). *Update rate* is the frequency with which data samples are measured and reported, usually reported as "samples/second". *Transport delay* is the amount of time that it takes data to travel through the system and become available for use. *Latency* (or *throughput*) usually refers to the amount of time required to accurately reflect a change in the quantity being measured. It is influenced by pure transport delay and also by dynamic operators (for example, a low pass filter) in the signal path. *Bandwidth* is the range of sinusoidal input frequencies that can be processed by the system without significant attenuation or distortion. A more detailed discussion of performance parameters can be found in Kocian and Task [1].

3.1 MECHANICAL HEAD TRACKING

Mechanical head trackers, sometimes referred to as goneometers, work by mechanically coupling head gear to the environment (e.g. airframe) through a set of linkages connected by flexible joints. The position of each joint is

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

measured by a transducer, and the set of joint positions is used to calculate head gear position and orientation in 6 degrees of freedom. Transducers are typically optical encoders, potentiometers, strain gauges, or some combination of these.

There are a very small number of commercially available mechanical devices which are specifically designed to track head gear position and orientation. Many "one of a kind" goneometers have been built for use in research and simulation laboratories. One such device, developed for use with a flight simulator [2] is sketched in Figure 1.

Some custom mechanical systems have been flight tested in various countries, especially on helicopters, and are usually designed to provide only azimuth and elevation degrees of freedom. For example, such a system has been used on the Cobra helicopter for many years. The system used on the Cobra [3] consists of an overhead slider mechanism allowing a rod to slide for and aft just above the pilots head. The rod is attached to the slide track with a universal joint and also has a universal joint on the other end which can be attached, via a nipple shaped magnet, to a mating receptacle on the pilot's helmet. The magnetic helmet attachment mechanism allows for very quick disconnect. The universal joint angles are measured with AC resolvers. Analog outputs from the resolvers are input to an electronics unit which computes the azimuth and elevation angle of the pilot's helmet (2 degrees of freedom), and sends a corresponding command signal to a rotating gun mount, or to a wire guided missile system.

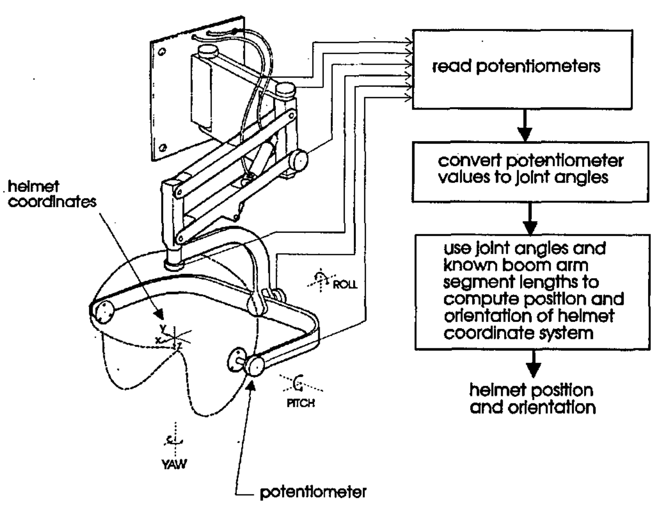


Figure 1. Sketch of mechanical head tracker built for use in a flight simulator, redrawn from Jarrett [2].

Mechanical trackers can have relatively low cost, and are capable of good accuracy, high update rate, reasonable range for a seated user, and very good dependability; but the mechanical linkage takes up valuable cockpit space, are subject to mechanical damage, are affected inertial forces, and pose a difficult ejection safety problem. In spite of excellent performance parameters, future in-flight use of mechanical head trackers is likely to be restricted to helicopter, transport, or ground based, applications, and then only when low cost is important.. Mechanical trackers will probably continue to be extremely useful as low cost research and development tools.

3.2 INERTIAL HEAD TRACKING

Inertial sensors are available which can measure angular velocity and specific force (the vector sum of gravity and acceleration forces) with respect to an inertially stable reference frame. Methods for position and orientation tracking with such instruments have been developed for inertial navigation and the same principles can be applied to tracking a person's head gear. If an initial orientation is known, angular velocity can be integrated to continually estimate orientation angle. Once orientation with respect to gravity is known, gravity can be subtracted from specific force data to yield acceleration with respect to the gravitational field. If an initial position and velocity are known, acceleration can then be integrated to continually estimate current position and orientation. Inertial sensors measure motion with respect to an inertially stable reference frame; so in order to measure head motion with respect to an aircraft cockpit, information from an inertial package that is fixed to the airframe must be subtracted from measurements made by the head mounted package

Transient errors in the angular velocity or acceleration measurements accumulate in the integrated orientation and position estimates. Even if the inertial components are quite accurate this "dead reckoning" technique requires periodic independent measures of position and orientation to remove accumulated drift. The rate of drift, and consequently the frequency with which it must be corrected, depend on the accuracy of the sensor measurements and of the integration process. With a sensor package that is of practical size and weight for head mounting, drifts of at least several degrees/minute and several cm/minute would not be unexpected

Inertial sensors provide high bandwidth angular velocity and acceleration information, and can provide position and orientation information with very high resolution, but the requirement for frequent drift correction constrains inertial head tracking to use in conjunction with other head tracking techniques. There is currently a commercially available system that uses a combination of acoustic and inertial sensors to measure head gear position and orientation. An early version of this device is described in Foxlin and Durlach [4]. The head mounted inertial package measures approximately 3.5 cm x 3 cm x 3 cm. The device was not, however, intended for use on an aircraft and the current system makes no provision for subtracting vehicle motion. Inertial sensors, particularly angular rate sensors, have been used quite successfully to add high frequency (lead) information to systems employing other head tracking techniques. For example, Emura and Tachi [5] describe an optimal estimation technique for combining inertial angular rate information with magnetic head tracker data.

3.3 ACOUSTIC HEAD TRACKING

Acoustic trackers use a triangulation technique that is usually based on sound propagation time. The ultrasonic frequency range is generally used so as not to be audible to people.

Assuming that the speed of sound is known, the delay between sound emission by a speaker, and detection by a microphone yields the distance between speaker and microphone. Note that this assumption can be compromised by changes in the speed of sound due to temperature changes or other atmospheric changes. Distance values from 3 known fixed receivers (microphones) to a moving speaker allows the emitter (speaker) position to be triangulated. The emitter is usually the moving component since a single emission from

one speaker can easily be received by multiple microphones without confusion..

Line of sight must always be maintained between the emitters and receivers since it is assumed that sound can follow a straight trajectory between emitter and receiver.

If at least 3 such speakers are fastened in known positions on a helmet, the helmet position and orientation can be unambiguously computed.

A small number of commercially available systems have been designed primarily for use as 3D computer input devices. A device was made in the 1980s to acoustically detect pilot head orientation (3 rotational degrees of freedom) for weapon aiming application, but is no longer available. A commercially available device mentioned in the previous section on inertial tracking [4], combines acoustic steady state measures with higher bandwidth inertial measures to implement a head tracking device. The resulting system is intended to have update and throughput rates as well as resolution characteristics (ability to measure small changes) that are associated with inertial systems, while maintaining the steady state performance characteristics of acoustic trackers.

It is also possible to detect motion of an emitter with respect to a receiver by measuring phase changes between a signal and reference sound source [6]. This has the same inherent problem as inertial sensing in that no steady state measurement is made; rather, a velocity measure must be integrated.

Acoustic trackers require line of sight between emitters and receivers, are easily influenced by temperature gradients and air currents, and are subject to interference from echoes and other acoustic sources, especially in the noisy environment of military aviation. Update rate is limited, primarily by the speed of sound, to about 30 samples/sec.

Currently available acoustic tracking devices are not as accurate or dependable as the state of the art magnetic or optical tracking devices, and militarized versions are not currently available. Acoustic devices do not suffer from metal and electro-magnetic interference as do magnetic systems, or from sunlight interference as do optical trackers; but the problems listed above are at least as severe. Future development of acoustic technologies may solve or reduce the practical problems, but at present both magnetic and optical technologies are significantly more mature and are more likely to find practical use in airborne environments.

3.4 OPTICAL HEAD TRACKING

Over the past 35 years engineers have developed a variety of optical helmet tracking systems in an attempt to attain a satisfactory balance between measurement accuracy and reliability in the cockpit environment. Although several have exploited phenomena such as interferometry and pattern recognition [7], the most successful have been based upon triangulation. These invariably use near infra-red light, which is unnoticeable to the user and for which a variety of commercial emitters and receivers are available, and they all measure a set of angles between cockpit- and helmet-mounted devices. They differ by employing alternative devices, and in some the emitters are fixed in the cockpit while in others they are on the helmet. Their sensitivity to artifacts, particularly those due to incident sunlight, also depends strongly on the chosen sensor.

The Honeywell MOVTAS (Modified Visual Target Acquisition Set), shown schematically in Figure 2, was devised in the late 60's and has been installed in a variety of

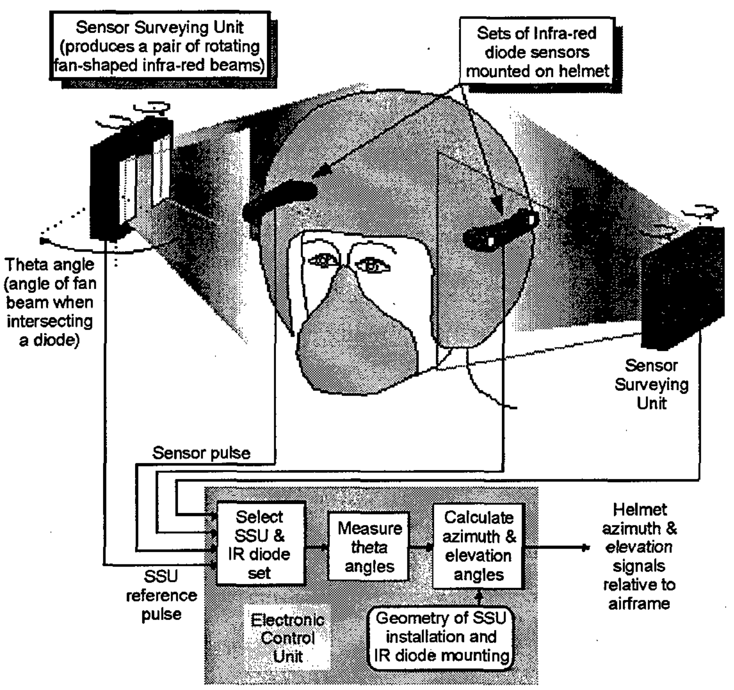


Figure 2. Schematic summarizing the Honeywell MOTVAS optical head tracker

aircraft. It is best known as the helmet tracker employed in the IHADSS (Integrated Helmet and Designating Sub-System) for the AH-64 Apache helicopter, in current US Army service. As illustrated, a helmet-mounted infra-red sensing diode produces a short electrical pulse when illuminated by a fan-shaped beam from a sensor surveying unit (SSU) mounted in the cockpit. The principle is analogous to a sailor observing the flash of a lighthouse. As the beam rotates rapidly at a constant angular rate, the interval between detection by the helmet-mounted diode and a reference pulse produced by the beam rotating mechanism is proportional to the mechanism angle at the instant the beam illuminates the diode. The diodes are paired, and a pair is "surveyed" by both beams in a SSU to give four beam angle measurements. Given knowledge of the installation dimensions, the electronic unit solves the trigonometric equations to calculate the helmet pointing direction, which is output each computational cycle as the helmet azimuth and an elevation angle. Several sets of diodes and SSUs are normally used to extend the range of measurement and the head box.

A more modern approach is illustrated in Figure 3. Here, a cluster of LED emitters on the helmet is imaged by a cockpit-mounted camera. An electronic unit, based on digital signal processing (DSP) chips, finds the position of each diode in the 2-dimensional camera image and, knowing the installation geometry and the distortion introduced by the camera optics, calculates both the position and the orientation of the helmet. The update rate of systems employing video cameras as imaging sensors is usually limited by the frame rate of the video signal to either 50 or 60 Hz, although fast frame cameras can be employed to increase the measurement frequency. Measurement delay can be reduced by motion prediction algorithms, and sensitivity to sunlight can be reduced significantly by only opening the camera electronic shutter during the brief fraction of the frame period when the diodes are pulsed.

Some systems use lateral effect photo-sensitive detectors (LEPSD) instead of video sensors [7] to increase the measurement update rate and enable sequential pulsing of individual diodes to remove any uncertainty in their identity and improve signal delectability. It is essential to filter the incident light to exclude all but the IR source waveband to prevent sunlight from saturating the detector, but it is possible to compensate for the in-band sunlight by sampling the LEPSD output when all the diodes are momentarily inactive.

As with the MOVTAS system, the range of measurements and the allowable head box of the imaging techniques are invariably extended using several clusters of emitters and several cameras. The allowable range of head positions has been taken further in a ground-based laboratory where the user can walk around a room in which the ceiling is studded with clusters of IR emitters [8].

Optical systems require very careful placement of cockpit mounted units to yield the required range of measurement, allow an adequate head motion envelope, and adequately shield the sensors from direct sunlight, all without intruding on the pilot's view through the canopy.

The helmet-mounted and cockpit-mounted units must be installed where they give the required range of measurement and an adequate head motion envelope without intruding on the pilot's view through the canopy. The sensors should also be shielded from direct sunlight, and the canopy should not reflect either the sun or the IR emissions into the sensor field.

At night, the mixture of emitted, reflected and scattered IR

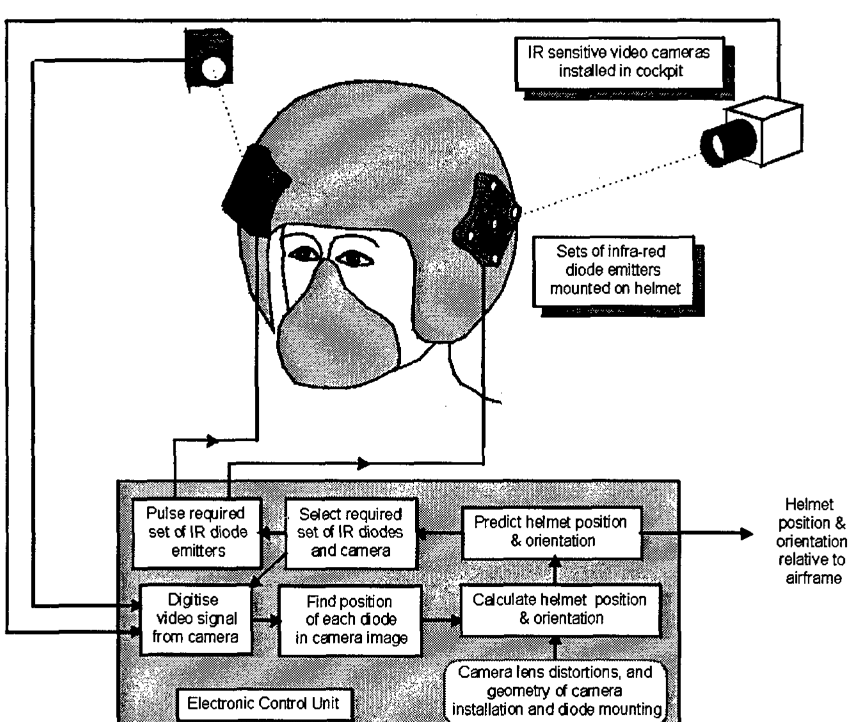


Figure 3. Schematic summarizing a modern optical head tracker

from the SSUs makes MOVTAS incompatible with the use of night vision goggles. A similar intensifier overloading can occur with the later optical trackers, particularly when helmet-mounted diode emissions are reflected from the canopy. There is also some concern that IR emission from the cockpit could make military aircraft more readily detected by external surveillance systems.

Although optical trackers offer good performance and require no calibration or alignment in service, they may be susceptible to strong sunlight during daytime and at night they may interfere with other cockpit systems which utilize the IR spectrum. Given that electro-magnetic tracking systems achieve comparable performance with none of these attendant drawbacks, and at similar cost, optical techniques are unlikely to be preferred.

It is possible that a simple optical tracker, working around a small cone of angles centered on the boresight, could be installed to complement an electro-magnetic system. The optical tracker could have the very high accuracy for delivering boresighted weapons, and it could alleviate the need for pre-take-off harmonization of the e-m system. Cross-checking would also ensure that the helmet tracker of a visually-coupled system was unlikely to produce erroneous, and potentially disorienting, measurements.

3.5 MAGNETIC HEAD TRACKING

Magnetic trackers create magnetic fields of known orientation and measure the current induced in sensor (receiver) coils that are fixed to the object being tracked.

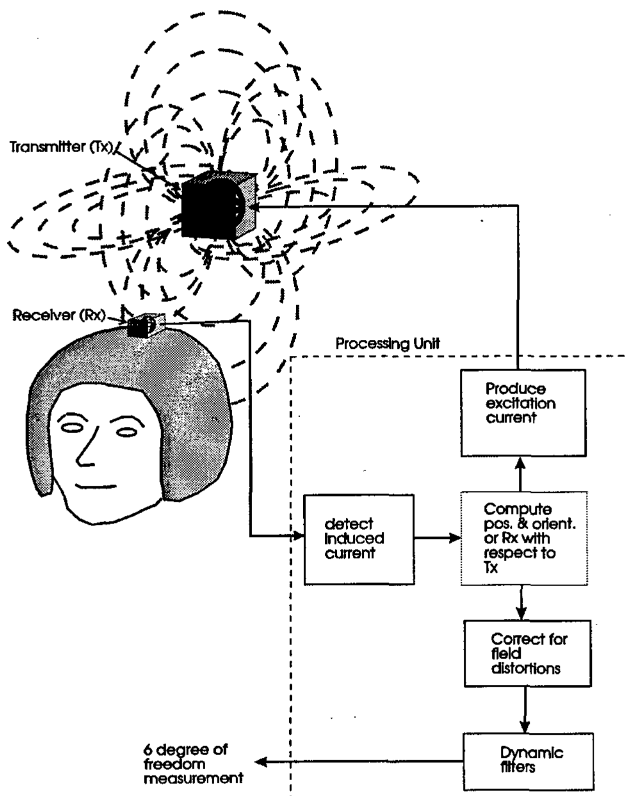


Figure 4. Schematic showing a generic electro-magnetic head tracking system.

As shown schematically in Figure 4, a set of 3 orthogonally oriented coils mounted to the environment (e.g., airframe) are sequentially excited with electric current, sequentially producing electro-magnetic fields with mutually orthogonal polarization's. This set of antennae is usually referred to as the transmitter or source.

A smaller set of orthogonal coils, usually referred to as the sensor or receiver, are mounted to the object being tracked (e.g., aircrew headgear). The current induced in each of the 3 sensor coils is measured during the field produced by each of the 3 transmitter coils. The 9 sensor responses are processed to compute position of the sensor with respect to the transmitter in 6 degrees of freedom [9,10].

The transmitter is typically housed in a cube shaped enclosure, ranging from 5.5 to 10 cm on each side. The sensor is typically housed in a much smaller enclosure, typically 1.5 to 2.5 cm on each side.

Two categories of magnetic system are available: those using an AC coupled technique and those using a DC technique. AC type systems excite each transmitter antenna with a sinusoid and can take advantage of AC coupling techniques to eliminate the effect of static fields in the environment. AC systems are very susceptible, however, to error due to the presence of conductive metal in the environment. The errors are due to eddy currents induced in the conductive metal by changing fields.

DC systems excite each transmitter antenna with a DC current pulse. Sensor antennae are sampled when the transmitter is dormant, as well as during the time each transmitter antenna is excited, so that components of the Earth's magnetic field can be subtracted. When run with update rates in the region of 100 Hz, DC systems are far less sensitive to the presence of conductive metals than are AC systems. The eddy currents produced by field changes die out at an exponential rate proportional to the metal conductivity. As update rates increase and there is less time during each transmitter antenna pulse to wait for eddy currents to die away, DC systems become more susceptible to eddy current interference [11,12].

This is now a relatively mature technology, and magnetic tracking devices of both AC and DC type are readily available in both commercial and militarized versions.

In a benign environment (no large metal objects or electro-magnetic interference problems), commercial type systems typically offer accuracy ranging from 0.75 to 2.5 mm translation, and 0.15-0.5° orientation. Accuracy is usually best when sensor and transmitter are very close, and tends to decrease as they separate. The allowable motion box is typically on the order of a 1 meter hemisphere for best performance. Update rate typically ranges from 60 -120 Hz, and latency ranges from 4-150 msec with a typical value of about 40 msec depending on the type of system and amount of filtering used.

Depending on the environment, varying amounts of filtering may be needed to reduce noise in the measurement. The filters used are usually dynamic filters with properties that are related to motion rates, and this makes latency determination very complex. A comparison of latency's in some commercially available systems can be found in [13].

A promising approach to reducing system lag, as described by Emura and Tachi [5] (and previously mentioned in the section describing inertial trackers), is to augment the magnetic system data with information from inertial sensors. The magnetic system provides accurate low frequency

information, while angular velocity sensors can provide very good high frequency information.

Metal objects produce errors whose magnitude depends on proximity to the magnetic components as well as size and composition of the metal object. It is possible to compensate for effect of stationary metal, but determination of compensation equation parameters is an elaborate procedure requiring placement of the sensor in many precisely known positions with a non metallic jig. The results are then valid only for one precisely defined physical environment. Such procedures, referred to as cockpit mapping, usually take several days to be completed with an acceptable accuracy. Successful transfer of mapping data from one aircraft to another of same type is possible only if very tight manufacturing tolerances are maintained.

Another problem has been posed by metal objects attached to the aircrew head gear, and subject to repositioning as helmet mounted systems are reconfigured for different tasks. This problem has been solved, or at least reduced to a manageable level by incorporating miniature compensating circuitry at the magnetic sensor [14].

Electromagnetic emissions from other equipment can also effect the magnetic field and cause error which usually manifests itself as high frequency measurement noise. This type of error can often be eliminated or reduced by properly synchronizing the magnetic system with the offending electro-magnetic source.

Current state of the art does allow magnetic head tracker problems to be managed successfully in most cases. It has been reported, for example, that a militarized AC magnetic tracker, developed to have a high degree of metal tolerance, has achieved angular accuracy's of 0.1° RMS, within mapped areas, even in environments containing a great deal of interfering metal. This performance has been achieved, for example, in an OH-58 helicopter cockpit for sensor motion within an 18" x 12" x 7" motion box [15].

Magnetic tracking technology is relatively mature, has been militarized, and offers the best overall head tracking

Table 1. Summary of Major Head Tracking Techniques

Method	Major Characteristics	Typical Performance	Status
Mechanical	• Good accuracy • High bandwidth • Low cost • Subject to inertial forces and mechanical damage • Takes up a lot of cockpit space • Mechanical linkage between helmet and cockpit is undesirable (ejection and fast egress problems)	• accuracy: ~5 mm; ~0.2° • update rate: >500 samples/sec • (can vary significantly with specific implementation)	• Has seen operational in-flight use in the past (usually on helicopters for 2 degree of freedom application) • Future use will probably emphasize ground based simulation, R&D, use on helicopters or transports when very low cost system needed.
Inertial	• High bandwidth • Poor static accuracy (requires time integration of accelerations and angular velocities)	• accuracy: ~0.1-1°/sec: ~0.002-0.2 m/sec^2 (not appropriate for static measurement) • update rate: >500 samples/sec	• Potential use in conjunction with other techniques that have good static accuracy.
Acoustic	• Moderate Accuracy • Moderate to poor bandwidth • Echo and blockage problems • Environment noise interference problems • Effected by air temperature and motion	• accuracy: ~5 mm; ~0.5° • update rate: ~30samples/sec	• Requires further work to match optical and magnetic system performance • Systems currently in production are intended primarily for ground based virtual reality applications. • A system is available commercially which combines acoustic and inertial techniques
Optical	• Good accuracy • Moderate to poor bandwidth • Stray IR interference problems (especially from sunlight) • IR emissions may interfere with other cockpit systems that use IR. • Camera mounting problems (multiple cameras must be properly positioned) • Line of sight interference problems	• accuracy: ~1 mm; ~0.2° • update rate: 30 samples/sec	• Mature technology • Military versions available (have seen operational use). • Currently under-perform magnetic systems at similar price
Magnetic	• Very good accuracy • Moderate bandwidth • Large motion box • Metal (including helmet mounted metal) interference and electromagnetic emission problems have largely been solved for most environments, but create expensive and time consuming installation and calibration requirements.	• accuracy: ~1 mm; ~0.1-0.2° • update rate: ~120 samples/sec	• Mature technology • Military versions available • In current operational use • Further accuracy improvement might enable implementation of head mounted HUD

performance available at this time. It is likely to be the predominant head tracking technique for the next generation of military head coupled systems.

All of the major head tracking techniques are summarized in Table 1.

3.6 REQUIRED DEGREES OF FREEDOM

Parallax is the difference in sighting angles necessary to sight the same object from different positions. If separation between two sighting positions is small compared to the distance of each from the target, parallax will be negligible (E.g., two telescopes aimed at the same star will be parallel to each other).

Designation of distant external targets by head pointing usually requires measurement of only head azimuth and elevation (2 rotational degrees of freedom). The parallax effect of motion within the cockpit is minimal in this case because head motions are small compared to the distance from the targets; and since roll is a rotation about the pointing axis, it doesn't affect the direction of the pointing vector. Stabilization of HMD imagery with respect to the external environment usually requires measurement of all 3 rotational degrees of freedom. Parallax is still not a significant effect, but the imagery must be stabilized in Roll as well as in pitch and yaw. Designation of objects within the cockpit or stabilization of imagery relative to the cockpit interior requires measurement of head position in all 6 degrees of freedom (3 position coordinates, and 3 rotation angles).

3.7 BORESIGHTING

When head position measurement is used to implement a head mounted aiming device, it is necessary to know the relation between the measured position of the head gear and the line of gaze ("boresight") produced when the pilot sights through a head mounted aiming reticule. The process of determining this relation is often referred to as "boresighting". It is usually accomplished with a calibration procedure during which the pilot, who is positioned so that his eye point is known, sights a target whose position is also known. The line of gaze is thus known independently of the head tracker measurement, and if the head tracker measurement is also sampled at this time, the two can be compared. If the eye point with respect to the head gear is precisely known, alternate procedures can be devised to accomplish the same result with appropriate jigs and laser beams so as not to involve the human pilot.

4. PHYSIOLOGICAL AND BEHAVIORAL CONSIDERATIONS

Normal range of head motion is approximately ±60° for chin up chin down (pitch) motion, ±40° for tilting one ear towards the shoulder, and just under ±80° degrees for rotation about the spinal column (yaw) [16,17]. All of these range values have standard deviations of close to 20% or more between subjects. For pilots head motion may sometimes be further restricted by flight gear.

Typical peak velocities for voluntary head motion are about 600 °/sec in yaw rotation and about half that for pitch, with virtually all frequency domain energy below 15 Hz [18].

Typical reaction to the appearance of a non predictable visual target is a rapid eye movement (saccade), followed by a head motion towards the target. The head movement typically begins 30-50 msec after initiation of the eye saccade. If the target appearance time and location is predictable, an anticipatory head motion typically precedes eye motion by up to several hundred msec [19, 20, 21, 22]. This typical behavior breaks down if the visual field is sufficiently restricted. Under these conditions, head motion alone may be used to direct gaze, and ability to perform visual tracking tasks is impaired [23].

There is no entirely natural way to point the head precisely. Designation of physical objects by head pointing alone (no eye tracking) requires a head mounted sighting reticule so that a fixed line of sight is defined with respect to the head. If head motion is used to control a display cursor, visible position of the cursor provides the necessary feedback.

Although turning the head toward a target is a natural action, neither fine positioning of the head nor maintaining rigid head positions for extended periods are at all natural.

Human performance for designating an eccentric target by sighting through a head mounted reticule can probably be best described by Fitts' law, which relates "time-to-target" to the ratio of (distance-to-target)/(target-size) [24, 25, 26, 27, 28, 29, 30, 31]. The farther the head must be turned to reach the target, and the smaller the size of the target, the harder the task and the longer it takes. The precise performance achieved is very dependent on task details as well as the performance of measurement and display equipment involved. One laboratory study, in a non-dynamic environment, showed that a particular head mounted sight implementation required 0.8-1.5 seconds to bring aim point to within 2.5° of a 0.2° diameter target, and 2-4 seconds to come within 0.3° of the target [32]. Once on target, also in a non dynamic environment, tracking with a head mounted sight has been demonstrated with RMS error of about 0.2° [33, 34, 35, 36]. The effect of variables such as helmet weight, reticule size and shape, and off boresight angle are reviewed in an article by Wells and Griffin [37].

Inertial forces have a detrimental affect on head motion and tracking performance; furthermore, the dynamics of these forces must be considered. High Gz levels make head motion more difficult. The head becomes noticeably heavy at 2-3 Gz and head motion becomes extremely difficult, if not impossible, at 8 Gz [38]. One set of centrifuge studies found that tracking error with a head mounted sight increased from 0.2° at normal gravity to 0.8-1° at constant 5 Gz levels. Changing acceleration ("jerk") caused even more error, averaging 1.5° and sometimes exceeding 5° during Gz onset rates of 1 Gz/sec [33, 34, 35, 36]. Resistance to sinusoidal force externally applied to the head has been shown to be nonlinear at some frequencies [39]. Head pointing is also significantly disturbed by whole body vibration, especially in the 3 to 6 Hz range (a little bit above the jolts transferred to the head of a runner). The predominant disturbance is an involuntary nodding of the head due to vertical (heave) seat motion. Sideways (sway) and fore/aft (shunt) motions have significantly less effect. Head pitching can be controlled voluntarily if the excitation is below about 0.5 Hz, while vibration above about 10 Hz is damped by the trunk [40, 41, 42, 43, 44].

Head worn mass and center of gravity location interact with dynamic forces to affect head mobility. It has been observed, for example, that when performing a smooth tracking task (tracking a visual target with head motion), under conditions of 5 Gz, some subjects cannot rotate their head beyond about 20° in azimuth and 40° in elevation [45]. No such limits are observed for ballistic head movements. It can be hypothesized that the effect is explained by changes in head center of gravity location with respect to neck pivot points.

5. APPLICATIONS

5.1 HELMET MOUNTED SIGHTS

The idea of providing the pilot of a combat aircraft with a helmet orientation sensing system and a simple monocular reticule display, so that he could designate an external target by moving his head to superimpose the reticule over the target, was devised in the early 1960s [7].

As shown in Figure 5, these two components formed a helmet-mounted sight (HMS) which was integrated into the weapon control system so that helmet orientation signals were sent directly to the seeker head of a lock-before-launch missile, such as the infra-red sensitive AIM-9L "Sidewinder", and the pilot would listen for the change in audible tone that told him when the missile had locked onto the target. He could then pull the trigger and release the missile. This was in contrast to the normal technique which required the pilot to use more extreme maneuvers to point the aircraft so that the target was brought within the small field of view of the HUD. Essentially the missile "launch success zone" expanded from a cone of about 10° to one of about 60° half-angle, which enabled him to exploit the inherent missile agility and attain earlier weapon release to win the combat.

Slight sophistication's brought further benefits. The signals from the helmet sensing system could also be used to point the aircraft radar so that the target range and range rate could be measured and the target g-level computed. Additional symbols in the reticule projector could then be used to tell the pilot whether the dynamically fluid relationship between the two aircraft represented a robust firing opportunity or merely a transitory chance shot. If the pilot looked away the radar would remain locked to the target, and arrow-shaped symbols alongside the projected aiming symbol could be illuminated to cue the direction in which he should move his head to re-acquire visual contact. A similar cueing arrangement could also help one crew member point out the target to another crew member.

Early equipment was developed by Honeywell in the form of the Visual Target Acquisition System (VTAS) which used a MOVTAS-type helmet tracker in conjunction with a simple robust reticule projector [8]. The pointing error arising from combined technological and human factors turned out to be comparable with the capture field of an infra-red missile, and the system was first deployed in a USAF squadron of F-4 aircraft.

Since then HMS systems have been developed by a number of manufacturers and the HMS has become an established facility in combat aircraft operated by the Air Forces of the US, Israel, SA and USSR. Systems are also likely to be retro-fitted to other fast jets such as Jaguar, Tornado, F-16 and F-15. The sight will be a standard requirement in all future combat aircraft such as Eurofighter-2000 and Rafale,

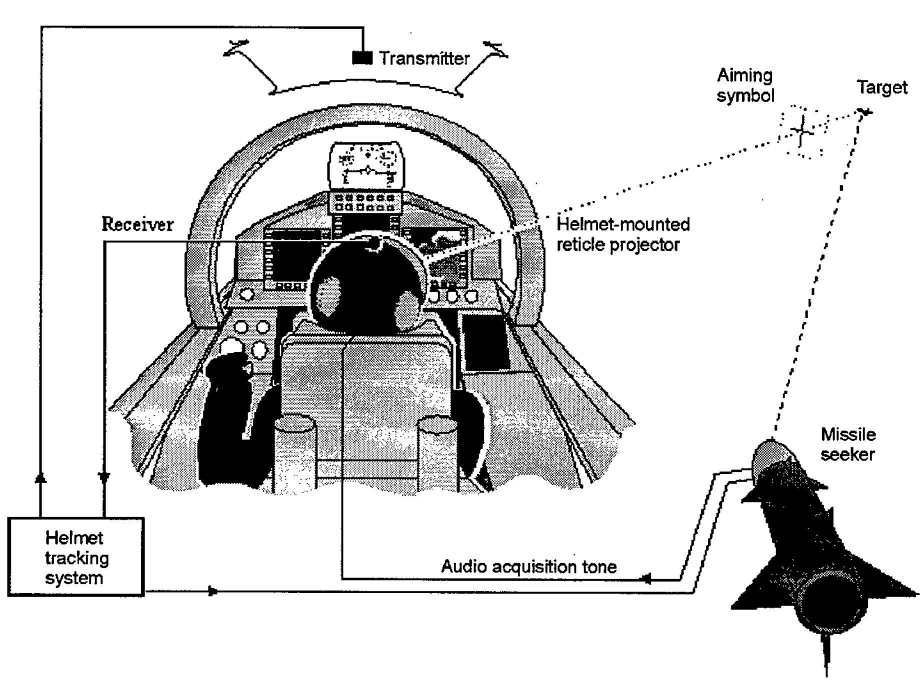


Figure 5. The basic elements of the helmet-mounted sight (HMS)

although in most of these aircraft the aiming symbol will be engineered as one element in a more complex set of imagery.

Note that for this application the helmet position sensing system need only measure the helmet line of sight relative to the airframe, which can be specified by two angles such as azimuth and elevation. Factors found to be most critical to usefulness of the HMS were:

- the brightness and sharpness of the reticule image,
- the size and positioning of the optical exit pupil,
- vibration-induced involuntary head motion,
- the difficulty of voluntary head motion at high-g,
- windscreen/canopy optical distortions, and
- the accuracy, update rate and head box size of the helmet tracking system.

5.2 VISUALLY COUPLED SYSTEMS

The idea of the visually coupled system (VCS), illustrated schematically in Figure 6, is a fairly obvious extension of the concept of the HMS to include the feedback of the image from a head-slaved sensor to a helmet-mounted picture-projecting display. Whereas the HMS is an explicit control, the visually coupled system concept adds implicit use of head position information.

When the field of view of the display matches that of the sensor, the user can have a reasonably normal visual sensation of viewing the world from the sensor location, although the resulting "synthetic vision" is likely to be somewhat limited in scope, quality and sharpness. In general, with suitable communication links and arrangement of the sensor, it is possible to give the user an ego-centric view from an inaccessible, hazardous or remote location, a facility which is currently under investigation for myriad applications ranging from micro-surgery to bomb disposal and tele-robotics

It is the use of a sensor, such as a thermal imager working in the atmospheric transmission spectrum between 8 μm and 14μm wavelength, which has been the most notable application. Such a VCS has been developed as the Passive Night Vision System (PNVS) to give the crew of the AH-64 Apache helicopter the means to fly at night and in conditions normally precluded by rain and fog, and not be blinded by missile rocket burn or gun muzzle flash [46].

The displayed sensor image is invariably overlaid by additional symbols giving flight and weapon aiming information, so the output of the helmet sensing system is simultaneously sent to the symbol generator. It is also available, via the avionics data bus, to the rest of the mission and weapon suite to enable the VCS to be used as a HMS. In daylight the system can operate exactly as the HMS described above, using an aiming cross in the center of the HMD field. At night or in poor visibility, when the sensor image is in use, the pilot can instead move his head so that the image of the target, rather than the directly viewed target, is designated. In the Apache it is also possible for the gunner/co-pilot in the front seat to receive a magnified target image from a narrow field of view sensor so that he can better identify and more accurately designate the target. However, since unwanted head shaking invariably disturbs his aim because head motion is also magnified, he also has recourse to a head-down

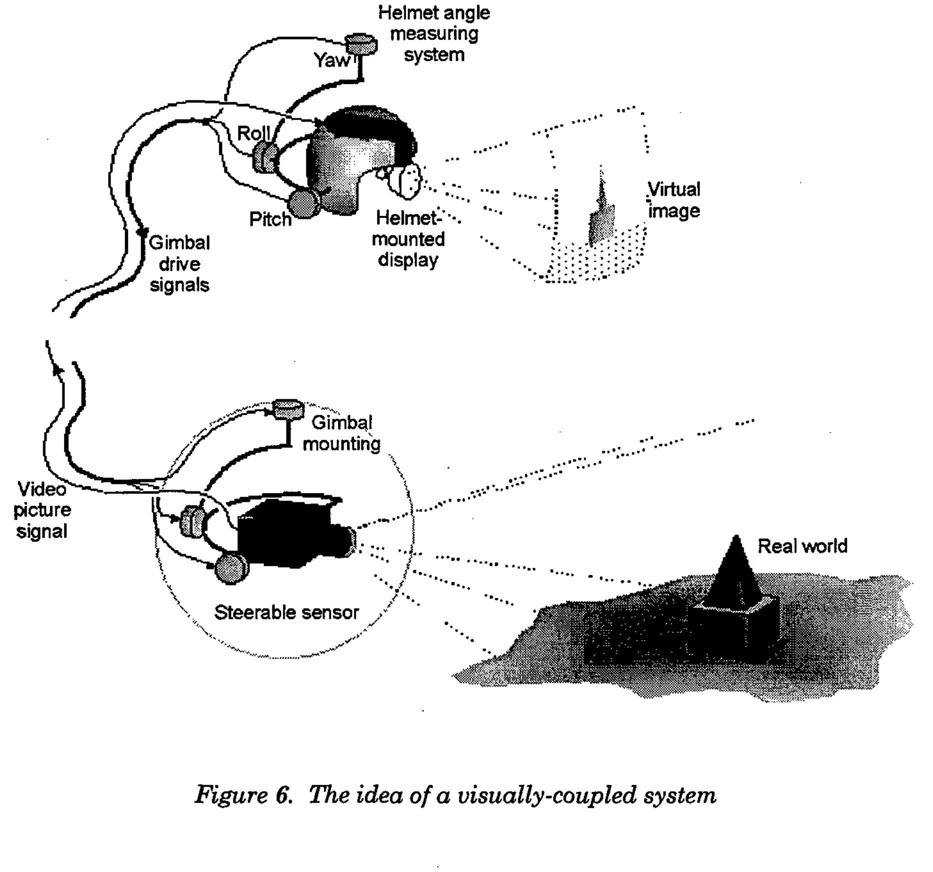


Figure 6. The idea of a visually-coupled system

display and a joystick to slew the sensor.

The technology employed in the PNVS is a monocular CRT display unit mounted on the side of the helmet, combined with a MOVTAS helmet sensing system. The next generation of helicopter VCS, such as those integrated into the RAH-66 Comanche and the Franco-German Tiger [47] will have binocular display systems and electro-magnetic helmet trackers. Similar equipment has been tested satisfactorily in fast jet trials aircraft [48, 49], and it is likely to be included in fixed wing combat aircraft which are soon to enter service, such as Eurofighter-2000. It is also under investigation as a means of supplying synthetic vision for future aircraft having windowless cockpits [50].

Note that for this application the helmet position sensing system must measure the helmet orientation in all three rotational degrees of freedom to give correct control over the sensor orientation.

5.3 HEAD UP DISPLAY

The requirement to replace an aircraft fixed Head Up Display (HUD) by presenting aircraft stabilized symbology within a helmet mounted display, and maintaining good registrational accuracy with the outside world, calls for head orientation measurement comparable with 1 milliradian alignment accuracy of current stationary HUD systems. The HUD application would require this accuracy only over a small forward cone of head pointing angles, and only for the set of HUD applications requiring accurate registration of symbology with real external objects (E.g., delivery of unguided bombs), but head tracking technology needs improvement to achieve this.

5.4 VIRTUAL COCKPIT

As summarized in Figure 7, the idea of the "virtual cockpit" (VC) is to extend the visually-coupled system to its practical limit so that it could provide an integrated and intuitive man-machine interface for all the tasks which make up the pilot's job [51, 52]. To enable operations in any external visibility condition, all relevant head-out information for controlling the aircraft, navigating, finding targets, avoiding threats and maintaining tactical awareness would be superimposed directly onto the pilot's normal view of the world or, when this is unavailable, the sensor-derived and computer-generated synthetic substitute for this view. Directional sound cues would provide reinforcement, and the stereoscopic capacity of the binocular display would allow the presentation of cockpit-stabilized 3-D "virtual panels" to convey aircraft systems information and tactical overviews. The idea also postulates that although the pilot would control the aircraft flight path and speed using conventional pedals, stick and throttle, and have ready access to HOTAS switches,

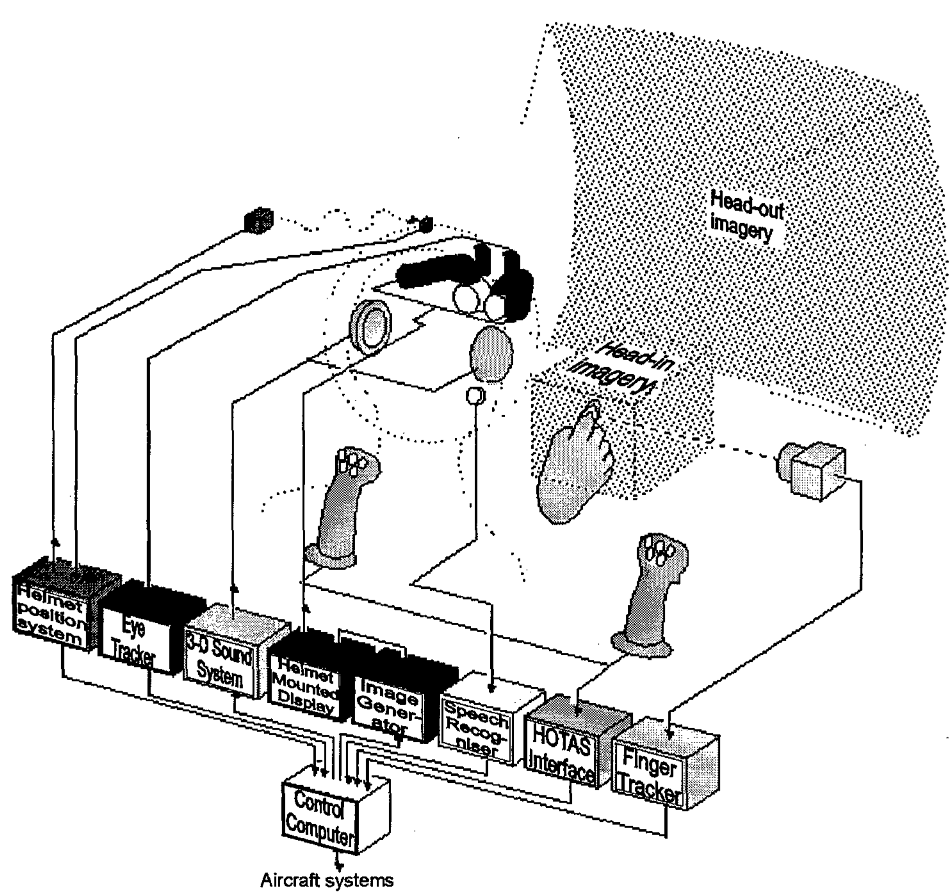


Figure 7. The likely systems of a virtual cockpit

he would also use a suite of novel controls which are compatible with virtual imagery.

Note that for this application the helmet position sensing system must measure all six degrees of motion freedom of the head relative to the airframe. Imagery must be stabilized with respect to the close interior surfaces of the cockpit and. parallax affects cannot be ignored

6. PROGNOSIS

Head tracking devices are a relatively mature technology compared to other enabling technologies for "alternative control" techniques.

Optical devices and both AC and DC type magnetic devices providing full six degree of freedom head position measurement are available in militarized configurations. These devices are in current use, although to a limited degree, in military aircraft.

Improvements are warranted to better handle potential interference conditions (e.g. sunlight for optical systems and moving metal for magnetic systems) and to provide better temporal response. In the case of magnetic systems the interference conditions can often be adequately handled but only with time consuming and expensive calibration procedures. Milliradian accuracy in operational environments would allow an expanded role for head tracking (E.g. head mounted HUD). Magnetic systems are making gains on this benchmark, but it has not yet been reliably achieved.

7. AKNOWLEDGEMENTS

This presentation is based on work done as part of AGARD working group 25, and contains contributions from all members of that group. It relies especially heavily on the efforts of two other members of that group: Dr. Don Jarrett, DERA, Farnborough, U.K,; and Dr. Alain Leger, Sextant Avionic, Bordeaux, France.

8. REFERENCES

1. Kocian, D. F., and Task, H. L., "Visually Coupled Systems Hardware and the Human Interface" In Barfield, W., and. Furness, T. A, (Eds) "Virtual Environments and Advanced Interface Design", New York, Oxford University Press, 1995.

2. Jarrett, D. N., "Helmet position sensor and loading mechanism", DRA working paper DRA-FS-93-WP892, 1993.

3. "Operator , Organizational, Direct Support and General Support Maintenance Manual", US Army Technical Manual TM 9-1270-212-14&P, July, 1981.

4. Foxlin, E.; and Durlach, N., "An inertial head-orientation tracker<with automatic drift compensation for use with HMD`s", in Singh, G., Feiner, S., K., and Thalmann, D. (Eds.) "Virtual Reality Software and Technology. Proceedings of the VRST '94 Conferrence", Singapore, 1994, pp 159-73.

5. Emura, A. and Tachi, S., "Compensation of time lag between actual and virtual spaces by multi-sensor integration", in "Proceedings of the 1994 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems", Las Vegas, NV, 1994.

6. Sutherland, I. E., "A head mounted 3 dimensional display", in "1968 Fall Joint Computer Conference, AFIPS Conference Proceedings", 33, 1968, pp 757-764.

7. Jacobs R. S., Triggs T. J., and Aldrich J. W., "Helmet-mounted display/sight system study" US Air Force Technical Report AFFDL-TR-70-83 Vol 1, 1970.

8. "An introduction to Honeywell helmet-mounted displays", Avionics Division, Honeywell, 1977.

9. Kuipers, J. B., "SPASYN -- an electromagnetic relative position and orientation tracking system", IEEE Transactions on Instrumentation and Measurement, IM-29, 4, 1980, pp 462-466.

10. Raab, F. H., Blood, E. B., Steiner, T. O., and Jones, H. R., "Magnetic Position and Orientation Tracking System", IEEE transactions on Aerospace and Electronic Systems, AES-15, 5, 1979.

11. Blood, E., "Device for Quantitatively Measuring the relative position and orientation of two bodies in the presence of metals utilizing DC magnetic fields", US patent no. 4,849692, 1989.

12. Blood, E., "Device for Quantitatively Measuring the relative position and orientation of two bodies in the presence of metals utilizing DC magnetic fields", US patent no. 4,945,305, 1990.

13. Adelstein, B. D., Johnston, E. R., and Ellis, S. R., "Dynamic Response of Electromagnetic spatial displacement trackers", Presence, 5, 3, 1996, pp 302-318.

14. Brindle, J. H., "Advanced helmet tracking technology developments for naval aviation", in "SAFE association, 33d Annual Symposium", Reno, NV, Oct. 23-25, Proceedings (A96-1671603-54), 1995, pp 34-53.

15. Hericks, J., Parise, M., and Wier, J., "Breaking down the barriers of cockpit metal in magnetic head tracking", in "Proceedings of the SPIE Head Mounted Displays", Orlando, FL, April 8-10, 1996.

16. Glanville, A. D., and Kreezer,G., "The maximum amplitude and velocity of joint movements in normal male human adults", Human Biology, 9, 1937, p 197.

17. Hertzberg, H. T. E., "Human Anthropology" in VanCott, H. P. and Kinkade, R. G., (Eds) "Human Engineering Guide to Equipment Design", American Institutes for Research, Wahington D.C., 1972.

18. Durlach, N. J. and Mavor, A. S., "Virtual Reality Scientific and Technological Challenges", Washington, D.C., National Academy Press, 1995, pp188-204.

19. Leigh, R. J., and Zee, D. S., "The Neurology of Eye Movements", Philadelphia, F. A Davis Company, 1983, pp 109-123.

20. Bizzi, E., "Eye-head coordination". In Brooks, V. B. (Ed) "Handbook of Physiology, The Nervous System", Sect 1, Vol 2, Part 2, Ch29, Bethesda, MD, American Physiological Society,. 1981, pp 1321-1336.

21. Bizzi, E., Kalil, R. E., and Morasso, P., "Two modes of active eye-head coordination in monkeys", Brain Research, 40, 1972, pp 45-48.

22. Mourant. R. R. and Grimson, C. G., "Predictive head-movements during automobile mirror sampling", Perceptual and Motor Skills, 44,. 1977, pp 283-286.

23. Sandor P. B., Leger A., "Tracking with a restricted field of view: performance and eye-head coordination aspects", Aviat. Space Environ. Med; 62, 11, 1991, pp 1026-31.

24. Fitts, P. M., "The information capacity of the human motor system in controlling the amplitude of movement", Journal of Experimental Psychology, 47, 1954, pp 381-391.

25. Welford, A. T., "The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress", Ergonomics, 3, 1960, pp 189-230.

26. MacKenzie, I. S., "Fitts' law as a research and design tool in human-computer interaction", Human Computer Interaction, 7, 1992, pp 91-139.

27. Jagacinski, R .J., and Monk, D. L., "Fitts' law in two dimensions with hand and head movements", Journal of Motor behavior, 17, 1985, pp 77-95.

28. Radwin, R. G., "A method for evaluating head-controlled computer input devices using Fitts' law", Human Factors, 32, 4, 1990, pp 423-438.

29. Lin, M. L., Radwin, R .G., and Vanderheiden, G .C., "Gain effects on performance using a head-controlled computer input device", Ergonomics, 35, 2, 1992, pp 159-175.

30. Borah, J., "Investigation of Eye and Head Controlled Cursor Positioning Techniques", US Air Force report AL/CF-SR-1995-0018, September 1995.

31. Spitz, G., "Target acquisition performance using a head mounted cursor control device and a stylus with digitizing tablet", in "Proceeings of the Human Factors Society 34th Annual Meeting", 1990, pp 405-409.

32. Barnes, G. R. and Sommerville, G. P., "Visual target aquisition and tracking performance using a helmet-mounted sight", Aviation, Space, and Environmental Medicine, April, 1978, pp 565-572.

33. Leger A., Sandor P., Clere J. M., Ossard G., "Mobilité de la tête et facteur de charge: approche expérimentale en centrifugeuse", AGARD-CP 471, AMP Symposium on "Neck injury in advanced military aircraft environments", Munich, Germany, 1989.

34. Leger A., Sandor P. "Désignation de cible sous facteur de charge: intérêt et limites du viseur de casque", AGARD-CP 478, AMP symposium on "Situational awareness in aerospace operations", Copenhagen, Denmark, .11, 2-5 October, 1989, pp 1-10.

35. Leger A., Sandor P., Troseille, X., "Désignation d'objectifs sous facteur de charge: poursuite de cibles mobile", R.E. N° 32 CEV/SE/LAMAS, 1990.

36. Leger A., Sandor P., Bourse C., Alain A., "Réponse biomécanique de la tête aux accélérations +Gz: Intérêt pour les études en simulation de combat", AGARD CP-517, "Helmet Mounted Displays and Night Vision Goggles", Pensacola, Fl, 6,-1991, pp 1-9.

37. Wells, M. J. and Griffin, M. J., "A review and investigation of aiming and tracking performance with head-mounted sights", IEEE Trans on Systems, Man and Cybernetics, SMC-17, 2, 1987, pp 210-221.

38. Fong, K. L., "Maximizing +Gz Tolerance in Pilots of High Performance Combat Aircraft", US Air Force Report AL-SR-1993-0001, December 1992.

39. Viviani P., Berthoz A., "Dynamics of the head-neck system in response to small perturbations: Analysis and modeling in the frequency domain", Biol. Cybernetics 19, 1975, pp 19-37.

40. Griffin, M. J., "Vertical vibration of seated subjects: Effects of posture, vibration level and frequency", Aviation, Space and Environmental Medicine, 46, 1975, pp 269-276.

41. Rowlands, G. F., "The transmission of vertical vibration to the head and shoulders of seated men", Royal Aircraft Establishment Technical Report TR-77068. Farnborough, England, 1977.

42. Lewis C. H., Griffin M. J., "Predicting the effect of vibration frequency and axis and seating conditions on the reading of numeric displays", Ergonomics, 23, 1980, pp 485-507.

43. Furness T. A., "The effect of whole body vibration on the perception of the helmet-mounted display", Ph. D. dissertation, Univ. Southampton (unpublished), 1981.

44. Tatham, N. O., "The effects of turbulence on helmet-mounted sight accuracies", AGARD CPP 267, 1979.

45. Leger A., unpublished observations, 1993.

46. Brown T. C., "AH-64 Apache night vision system" in "Night vision '92" Conference, London, 1992.

47. "Glass cockpit operational effectiveness", AGARD AR-349, 1996.

48. Church T. O. and Bennett W. S., "System automation and pilot-vehicle-interface for unconstrained low-altitude night attack" in "Combat automation for airborne weapon systems: Man-machine interface trends and technologies", Edinburgh, UK. AGARD-CP-520, 1992.

49. Lydick L. N., "Head-steered sensor flight test results and implications" "Combat automation for airborne weapon systems: Man-machine interface trends and technologies", Edinburgh, UK. AGARD-CP-520, 1992.

50. Rolwes M. S., "Design and flight testing of an electronic visibility system" in "Helmet-mounted displays II", SPIE 1290, 1990, pp 108-119.

51. Furness T. A. and Kocian D. F., "Putting humans in virtual space" The Society for Computer Simulation, Simulation Series, 16, 2, San Diego, CA, 1986, pp 214-230.

52. Kaye M. G., Ineson J., Jarrett D. N. and Wickham G., "Evaluation of virtual cockpit concepts during simulated missions", in "Helmet-mounted displays II" SPIE 1290, 1990, pp 236-245.

THE TECHNOLOGY AND APPLICATIONS OF BIOPOTENTIAL-BASED CONTROL

Grant R. McMillan
Air Force Research Laboratory (AFRL/HECP)
2255 H Street
Wright-Patterson Air Force Base, OH, USA 45433-7022

SUMMARY

This lecture reviews the technology for using electrical signals from the muscles and brain as a means for interacting with computers and other physical devices. It discusses the rationale for biopotential-based control technology, methods for acquiring and processing such signals from human operators, applications of these control technologies, and anticipated future developments.

1. INTRODUCTION

Electrical potentials can be measured from the natural electrochemical activity of many physiological systems (biopotentials). These signals are produced when excitable cells, such as muscle or nerve cells, are stimulated in response to an internal or external stimulus. We are interested in two types of biopotentials, the electromyographic (EMG) signals associated with the contraction of skeletal muscle and the electroencephalographic (EEG) signals associated with brain activity. As a control modality, the principal objective is to measure biopotential activity from the operator so that it can designate desired control actions or augment other control modalities.

2. THE RATIONALE FOR BIOPOTENTIAL-BASED CONTROL

It has been a goal of control system designers in many fields to tap our natural physiological systems to achieve intuitive, non-fatiguing control of external devices. The idea of an operator using natural motions of their hand to teleoperate a dextrous robot is one example where this intuitive mapping could reduce operator training and workload. Similarly, the notion of operating a device simply by thinking about the desired action represents the ultimate in intuitive control. Although current technology limits our ability to achieve such natural control systems, many practical devices have been designed and other promising technologies are being evaluated in the research community. For example, EMG-controlled prosthetic hands and wrists are of significant value for people with lower-arm amputations and thousands of units have been fitted worldwide (Figure 1). This area represents the most significant real-world application of biopotential-based control. This base of experience is reflected in the discussion of EMG systems, below.

In addition to applications as an assistive technology for persons with physical disabilities, biopotential-based control has a variety of potential applications in aerospace environments. These environments fall into two broad classes: (1) ones in which there are constraints on control access, and (2) ones in which there are high manual workload demands. An environment that requires operators to wear protective gear against chemical and biological agents is an example of the first class. The bulky clothing and gloves make it difficult, if not impossible, to operate small switches and controls. Extravehicular operation in space is another example. In addition to the limitations of the space suit, operators are constrained by the need to control the acceleration of their body when using tools and other devices. A third example is high acceleration flight in which g-forces essentially limit the pilots access to all controls except the joystick and throttle. The Hands-On Throttle and Stick (HOTAS) system is, in part, a response to the movement limitations of high acceleration flight.

A variety of aerospace applications fall into the second class, ones in which there are high manual workload demands. Maintenance technicians must devote high visual and manual attention to the task at hand. Frequent access to technical reference material is also required. Head-mounted displays and wearable computers are being developed to provide this information. However, the technician needs some means to interact with the information system, while keeping their hands devoted to their work. Voice control provides one option, but it can be constrained by high noise, the requirement for concurrent communication, as well as a variety of speaker characteristics. Biopotential-based control provides another option for such systems.

The control of secondary systems in flight can generate high manual workload for pilots. Although it is difficult to imagine a day when pilots might use biopotentials as a primary control, it is easier to foresee the use of biopotentials as a secondary control by which pilots or navigators perform multifunction display operation, weapons selection, radio frequency switching, or target selection.

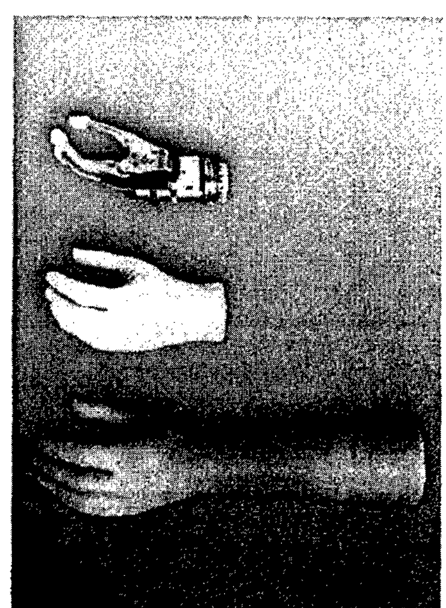


Figure 1. EMG-controlled prosthetic hand and arm systems. (Courtesy of Otto Bock USA, Minneapolis, Minnesota).

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

Operator state monitoring represents a third class of potential applications, but one that does not involve explicit system control. In this case the biopotentials provide on-line data, not otherwise available, about the operator's physical and cognitive state. Research in this area has emphasised three domains of human-system interaction that are of strategic importance to aerospace operations: operator workload monitoring, error prediction, and physical and cognitive fatigue monitoring. In addition to passive operator state monitoring, several advanced interface programs have considered the use of operator state data as part of an interface adaptation scheme. If used in this manner, biopotentials would provide an implicit system control function that blurs the distinction between monitoring and control.

3. CHARACTERISTICS OF EMG AND EEG BIOPOTENTIALS

3.1. EMG

The EMG signal resembles random noise that is amplitude modulated by changes in muscle activity (Figure 2). It results from the asynchronous firing of hundreds of groups of muscle fibres. The number of groups and their firing frequency controls the force produced by the muscle contraction [1, 2]. A convenient means of observing myoelectric activity is by an EMG recording on the surface of the skin. Surface-recorded EMG signals occupy the 20-500 Hz frequency band and are in the hundreds of microvolts to tens of millivolts amplitude range. Both of these characteristics present practical recording problems. The peak of EMG signal power is close to the power line frequency and the EMG amplitude is far less than the electrical interference due to capacitive coupling between the body and power mains.

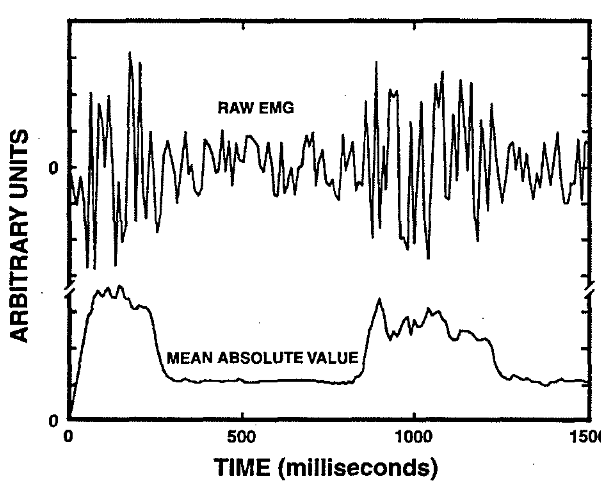


Figure 2. Time history of the raw EMG signal produced by two brief muscle contractions and the same signal after rectification and smoothing with a 100-millisecond moving average filter.

3.2. EEG

EEG recorded from the surface of the scalp represents a summation of the electrical activity of the brain (Figure 3). Although much of the EEG appears to be noise-like, it does contain specific rhythms and patterns that represent the synchronised activity of large groups of neurones. A large body of research has shown that these patterns are meaningful indicators of human sensory processing, cognitive activity and motor control. In addition, numerous EEG patterns can be brought under conscious voluntary control with appropriate training and feedback. The EEG signals of interest are in the 1-40 Hz frequency range with amplitudes ranging from 1-50 microvolts. Because of their small size, EEG signals are highly susceptible to contamination from eye and muscle activity, from external electrical sources and from movement of the user. These challenges can be managed, even in flight environments, but they require significant care and expertise on the part of system designers and operators.

4. THE TECHNOLOGY FOR ACQUIRING AND PROCESSING EMG AND EEG BIOPOTENTIALS FOR CONTROL

4.1. EMG AND EEG SIGNAL ACQUISITION

Although implanted electrodes continue to be explored for some biopotential control applications, it is unlikely that they will be employed in near-term aerospace environments. EMG and EEG signal acquisition is most commonly accomplished using metal, coated plastic or gel electrodes located on the surface of the skin. Mild cleaning of the skin is often performed to reduce the impedance of the electrode-skin interface. EMG electrodes are usually applied dry and rely on high input impedance amplifiers and the development of a perspiration layer to reduce common mode interference. EEG electrodes are commonly applied with a conductive paste or cream and affixed with adhesive rings, tape or an elastic band. Gel electrodes do not require a conductive paste since the gel itself contains an electrolyte. Aerospace applications will benefit from convenient dry electrode systems, but these are not yet commercially available for EEG recording. Bandpass and notch filters are commonly employed to eliminate DC drift, AC line noise and to focus on the signal frequency range of interest. Commercially available biological signal amplifiers are well suited to the amplification and filtering of both EMG and EEG signals.

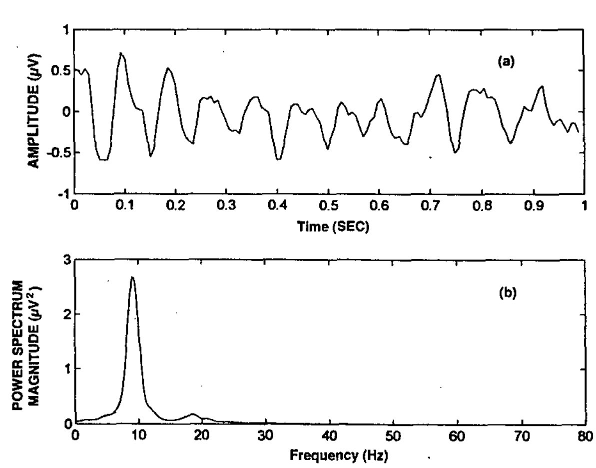


Figure 3. Sample raw EEG signal (a) and power spectral density (b). Spectrum based upon 10 seconds of data from one subject with 1 second of raw EEG shown. High-pass filter set to 1 Hz and low-pass filter set to 40 Hz. Spectral data were smoothed using Hanning window techniques. Subject's eyes were closed producing a marked increase in power in the alpha (10 Hz) region, visible in both plots.

Biological instrumentation amplifiers commonly use a differential arrangement to reduce the power line interference, but the interference rejection is only effective when the electrodes and skin are in contact and when the electrical impedance of the skin is low (Figure 4). Another problem associated with surface electrodes is that if they move relative to the skin surface, a noise signal is produced which can be confused with the true biological signals. In severe cases, this motion artefact completely overwhelms the biopotential and appears to the system as a large control signal. To minimise this unwanted effect a good electrode/skin contact must be maintained.

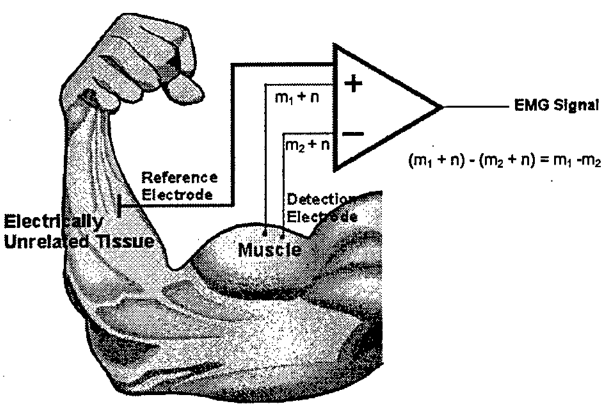


Figure 4. Differential signal recording arrangement commonly used for biopotential signal acquisition. The EMG signal is represented by m and the noise by n. (Courtesy of Carlo J. De Luca, Boston University, Neuromuscular Research Center - http://nmrc.bu.edu/nmrc/detect/emg.htm).

The next signal acquisition step in most biopotential controllers is analogue to digital signal conversion. Signal processing is most commonly performed in the digital domain. Personal computer systems, in some cases with digital signal processing boards added, provide sufficient computational power to implement the signal processing approaches reviewed below. Thus the size, cost and weight of biopotential-based control systems are not serious constraints.

Two general approaches characterise most EMG- and EEG-based control systems:

- The use of EMG and EEG responses, not normally associated with motor control, to operate external devices. For example, learned control (self-regulation) of the EEG activity in a specific frequency band might be used to turn a switch on or off.
- The use of natural EMG and EEG patterns, normally associated with sensory or motor activity, to produce a similar response in an external device. For example, the remaining movement-related myoelectric activity in the arm of an amputee might be used to operate a prosthetic hand.

When using either type of biopotential signal, designers face a significant trade-off between achieving short response times and smooth control outputs. Several factors contribute to this problem. First, the biopotential pattern being used for control is typically a small component of the overall signal and must be discriminated from normal background activity. This signal processing takes time and can degrade the system response. Second, various signal filtering schemes are commonly used to eliminate the sources of artefact listed above. These filtering steps smooth the control output, but can also introduce lag or delay. Finally, approaches that require the user to voluntarily modulate or produce specific patterns have an additional source of variability that must be managed. While users are able to rapidly raise a specific EMG or EEG component above a set threshold, holding it in a stable state is difficult. The raw signal sometimes shows brief drops below threshold that must be managed by the control algorithm. The required signal averaging or smoothing adds additional lag or delay to the system. These limitations, while severe, are not unique to biopotential-based control. Most eye-gaze-based controllers face many of the same problems in discriminating intentional from spontaneous eye movements.

4.2. EMG SIGNAL PROCESSING

4.2.1. Processing Learned EMG Responses as Control Signals

To enable the EMG signal to be used as a means of control, some feature of the signal must be extracted and an association must be made between values of this feature and the desired control response. The simplest EMG feature that can be extracted is signal amplitude. However, due to the random nature of the underlying myoelectric signal generation process, the average value of the EMG signal is zero. Consequently, any attempt to filter the EMG signal to produce a smooth output for control purposes will result in a zero signal. To remedy this problem the signal must be processed to produce a signal that reflects the variance of the EMG signal. Although a square law device has been shown to be the optimum processor based on error probability [3], most controllers approximate this non-linearity using a full-wave rectifier. By amplifying, rectifying and filtering the EMG, a control signal can be obtained based on the effort of the voluntary muscle activity. Several types of control algorithms have been developed that employ this signal amplitude feature. These can be divided into three general categories: (a) Level coding, (b) Rate coding and (c) Pulse coding.

4.2.1.1. Level Coding

To control a single degree-of-freedom device the control signal can be derived from either the EMG signal level of a single muscle or from the two EMG signal levels from an agonist/antagonist muscle pair, i.e., flexor/extensor. In the first case, the dynamic range of the signal (the range between the noise and the maximum EMG signal produced) is divided into three regions by two switching thresholds giving a 3-way switch to control the state of the terminal device, e.g., off, hand open, hand close (Figure 5a). In the latter case, each signal controls one state switch, i.e., flexor EMG for hand open, extensor EMG for hand close. To avoid the situation where both switches are in the on position, i.e., co-contraction of the two muscles, control is given to the larger of the two signals or to whichever signal first exceeded the switching threshold.

4.2.1.2. Rate Coding

Rate coding works on the principle of how fast the user contracts the control muscle (Figure 5b). A control signal is derived based on the initial slope of the processed EMG signal from a single electrode site. A slope threshold is set such that a slow contraction selects one function, i.e., hand open, and a fast contraction selects another, i.e., hand close. The operator performance and training requirements are similar to the single channel level-coded system.

4.2.1.3. Pulse Coding

It is also possible to derive a control signal based on pulses of EMG activity (Figure 5c). A simple coding scheme can be devised to define the control output signal. Function selection is then just a matter of producing the associated pulse code, i.e., one pulse - hand open, two pulses - hand close.

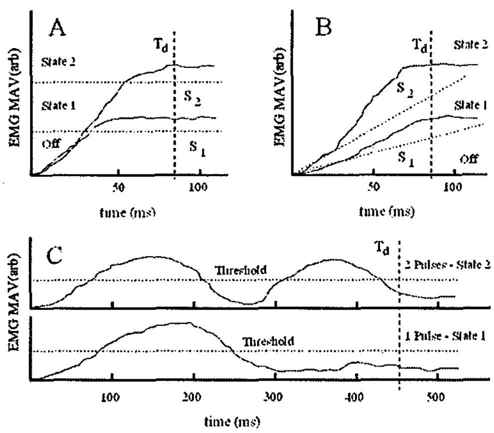


Figure 5. EMG signal coding: (a) Level, (b) Rate, (c) Pulse. MAV = Mean Absolute Value with arbitrary units.

4.2.1.4. Discussion

Clinical experience has shown that control systems based on either level coding or rate coding are easy to operate and that operator error is insignificant (Figure 6a) after a short period of training [4]. However, if the dynamic range is segmented into more than three regions (Figure 6b), in an effort to extract more control information, the operator error increases quickly [5]. Gains and switching level settings for each system depend on the individual's EMG levels and must be adjusted to achieve optimum control. Further adjustments may be required during the initial training period but little adjustment is required thereafter.

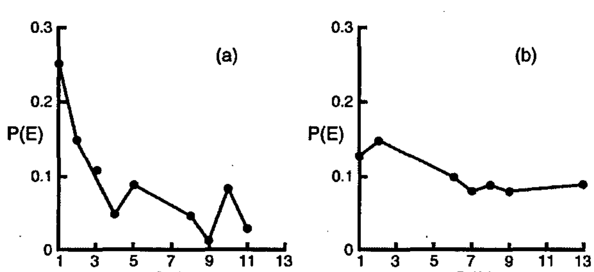


Figure 6. Training effect on the probability of operator error, P(E), for (a) 3-state and (b) 5-state level-coded EMG control systems. From [4].

For both level-coded and rate-coded systems the operator does not notice the small time delay introduced by the control system. A system based on pulse coding, however, introduces a noticeable delay due to the operator's inability to produce rapid EMG pulses. Although the delay limits the application of pulse coding to low bandwidth operations, this form of coding does have the potential to control a large number of functions.

4.2.2. Processing Natural EMG Responses as Control Signals

Several recent approaches to EMG control are based upon the interpretation of spontaneous EMG signals associated with natural muscle contractions. The impetus for such systems is that they would require little or no user training. No longer is the operator required to produce somewhat unnatural, self-regulated contractions. The control system learns to recognise the spatial and temporal patterns within the EMG signals, from one or several muscles, during contractions that correspond naturally to the desired controlled function. For example, an above-elbow amputee may choose to train the control system to associate the patterns produced during stump rotation with selection of wrist rotation. In other words, the training function is shifted from the operator to the control system.

4.2.2.1. Pattern Recognition

All myoelectric control systems implemented using pattern recognition have been based on the assumption that at a given electrode location, the set of parameters describing the EMG will be repeatable for a given state of muscle activation and furthermore that it will be different from one state of activation to another. To control m distinct functions requires m unique patterns of activity. Control schemes have been based almost entirely on the discriminant approach to pattern recognition, in which each pattern is described by a set of signal features. These features may be EMG from a number of myoelectric channels, a set of statistics describing the signal sampled at one electrode site, or some other reproducible set of features. Once the patterns are described in this feature space, an unknown pattern can be compared with them to determine which of the m functions should be selected.

In the most straight-forward approach, the activity (simply muscle active = on / muscle inactive = off) at a number of muscle sites is monitored and function activation is controlled on the basis of a match between the observed activity and a predefined on/off pattern across all sites.

4.2.2.2. Neural Networks

Much of the most recent work employs neural networks to classify specific patterns in the myoelectric signal from natural voluntary contractions of the residual limb [6]. In this case the pattern classifier is trained to recognise the specific contractions based on a set of time domain features. The features are extracted from a single EMG signal during reproductions of several contraction types. The classifier then uses this information to develop a feature template or signature for each contraction type (Figure 7). During use, each contraction produced by the operator is compared to all templates to determine which is most similar. The control system then selects the function that corresponds to this choice.

4.2.2.3. Discussion

The key advantage of a control system based on pattern recognition is that the training burden is moved from the operator to the control system. This assumes, however, that the operator will produce patterns that are unambiguous to the classifier. It is often the case that only a small number of distinct patterns can be found. Each new pattern class entered into the training set reduces the available feature space and increases the chance of pattern class overlap.

Pattern recognition systems based on on/off muscle activity patterns require many channels of EMG amplification and signal conditioning and a large number of accessible muscle sites. This is reduced in the single- and dual-channel systems, however, more complex signal processing hardware and software is necessary to achieve comparable system performance. The recognition of on/off events can also be

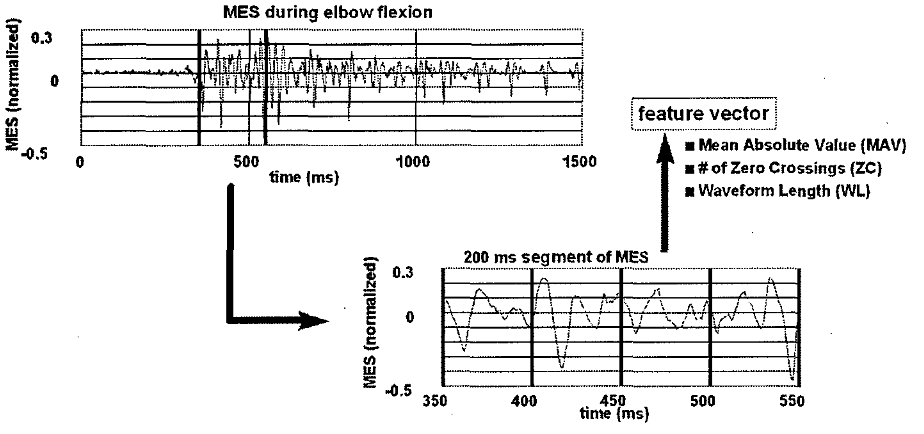


Figure 7. Representation of the EMG pattern by a time feature vector. MES = Myoelectric Signal.

done very quickly. Systems based on the recognition of more complex time or frequency domain features require an analysis of a much longer sample of the EMG signal to reduce the feature estimation error. This can introduce a noticeable time delay in the selection process.

4.3. USER FEEDBACK REQUIREMENTS WITH EMG-BASED CONTROL

Although EMG-based control provides muscle contraction feedback for users with intact sensory systems, many other feedback channels are absent. Nevertheless, most current prosthetic systems rely on visual feedback, or auditory and vibration cues from prosthetic motors, to provide this information. Attempts to provide grip force feedback in prosthetic devices have most often employed vibratory or electrical cues proportional to grip force. A study by Patterson and Katz [7] showed that pressure cues provided by an inflatable cuff permitted better grip force control than vibratory cues. However, visual cues alone appeared to be sufficient (Figure 8). To some extent this finding reflects limitations in the performance of current prosthetic devices, and it is generally believed that enhanced feedback will be required as the performance of prosthetic devices improve [8].

4.4. EMG-BASED CONTROL APPLICATION EXAMPLES

Early work sponsored by the US Air Force [9] showed that on/off EMG patterns from six sites on the upper arm could discriminate the purposeful muscular actions of pilots in simulated high-g environments. An EMG-based control system was designed which controlled the movement of a splint to provide a powered assist to the pilot's arm. The pilot was able to achieve 90% accuracy in a tracking exercise using this system.

The National Aeronautics and Space Agency (NASA) has sponsored several studies on the possibility of using EMG control for robotic teleoperation applications. Clark and Phillips [10] found that EMG time histories from normal hand and arm motion were not appropriate for controlling the complex movement kinematics of a robot arm. However, more recent work by Farry et al. [11] has found that a time-frequency analysis of the EMG patterns from forearm musculature could discriminate several different hand grasp types and thumb motions with a high degree of accuracy. Fernandez et al. [12] has continued this work and has used a classification scheme based on genetic programming to achieve 100% classification of thumb motions from the same EMG data. These results suggest that it may be feasible to use EMG from an operator's own hand and arm to replace or augment joysticks and exoskeletal instrumentation, and as the master to intuitively control a remote anthropometric robot arm.

Recent work by Junker, Berg, Schneider and McMillan [13] has shown that subjects can use a combination of EMG and EEG, referred to as a brain-body signal, extracted from electrodes on the forehead to control the movement of a cursor to track computer-generated targets. This group has also found that, for discrete on/off responses, a brain-body actuated control scheme can achieve high classification accuracy with little user training and with reaction times comparable to manual switches [14]. Vodovnik [15] has shown that reaction time can be enhanced using an EMG trigger. That study showed a substantial reduction in reaction time when an electronic braking system, triggered by EMG from the frontalis muscle, was used to augment a normal foot-activated automobile brake.

The EMG signal has the potential to augment more traditional control methodologies. For example, recent work [16] has shown that phonetically-relevant orofacial motions can be estimated from the underlying EMG activity.

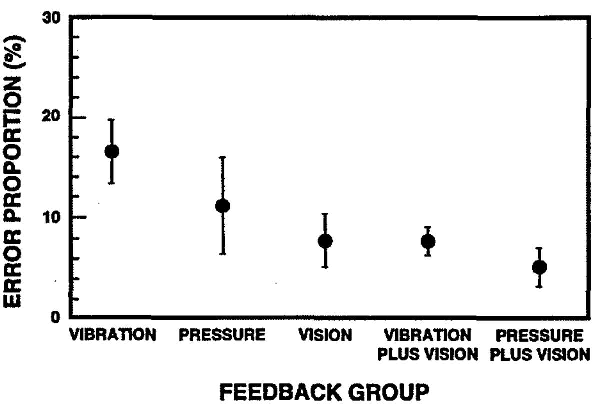


Figure 8. Grip force error with a prosthetic hand as a function of the type of feedback. Error magnitude is shown as a proportion of the reference force, i.e., force error/reference force. From [7].

Follow-on studies at the University of New Brunswick in Canada suggest that information from the EMG of facial muscles could improve the performance of current speech recognition systems (Figure 9). There is also a possibility that information from neck and shoulder muscle EMG could aid in determining head position and orientation. Kang et al. [17] have reported a 86.7% success rate in classifying ten head and shoulder movements using EMG pattern information from the Trapezius and Sternocleidomastoid muscles.

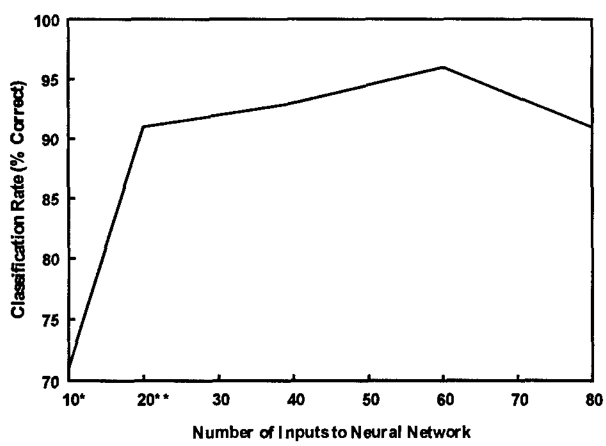


*Figure 9. Word recognition accuracy using EMG signals recorded from the face. Five electrode pairs were mounted in a flight oxygen mask and EMG data were recorded while one subject uttered the digits "zero" to "nine". Data were recorded for ten replications of each spoken digit. Inputs to the neural network were the ten pairwise ratios of the five EMG channels for time segment lengths ranging from one per word (10 inputs) to eight per word (80 inputs). * 6 residual errors in training set, ** 2 residual errors in training set.*

4.5. EEG SIGNAL PROCESSING

In most cases, the approaches used for EEG signal processing have been linked to a specific control application. Therefore, the signal processing techniques and associated applications are described in the same sections, rather than separating them as was done with the EMG discussion, above.

4.5.1. Processing Learned EEG Responses as Control Signals

4.5.1.1. EEG rhythm level coding

Level-coding techniques have been employed in several examples of EEG-based control. The EEG amplitude in a specific frequency band is determined using fast Fourier analysis, bandpass filtering, or some other technique, and this amplitude is compared to set threshold criteria or used as the input variable in a linear equation. For example, small amplitudes might move a computer cursor downward, medium amplitudes produce no motion, and large amplitudes might move the cursor upward.

Wolpaw and his colleagues [18, 19] developed such a system using self-regulation of the 8-12 Hz mu rhythm (Figure 10). Although it is in the same frequency range as the alpha rhythm, mu is recorded over the primary sensorimotor area of the brain and responds in known ways during movement preparation. In the single-axis task, the user moved the cursor to contact targets that appeared randomly at the top or bottom of the monitor. After approximately 18 hours of training, users required 2-6 seconds to move the cursor to a target. The target was correctly selected on 80-95 percent of the trials. Their dual-axis task used mu rhythm signals from both cortical hemispheres in a more complex control algorithm. The sum of the signals from the two hemispheres was used to control vertical cursor motion, while their difference controlled horizontal movement. After approximately 12 additional hours of training, the users required 2-4 seconds to move the cursor to targets that appeared in one of the four corners of the screen. The target was correctly selected on 40-70 percent of the trials.

4.5.1.2. Evoked Response Level Coding

Level coding of the amplitude of externally-evoked, as opposed to internally-generated, EEG signals has also been successfully employed. In this case, the brain response is produced by an external stimulus, such as a flickering light. With biofeedback and training, users can learn to modulate the amplitude of the brain's response to such stimuli.

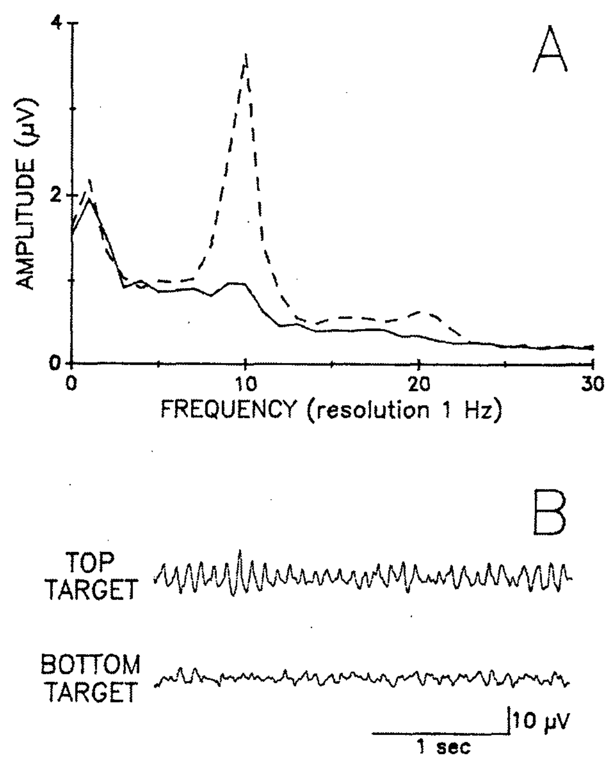


Figure 10. EEG rhythm level coding. High power in the 10 Hz region moves a cursor toward the top target and low power moves the cursor toward the bottom target. (a) Frequency spectrum of EEG signal for top (dashed line) and bottom (solid line) targets. (b) Sample EEG traces. From [19].

Using self-regulation of the visual evoked response, McMillan and Calhoun investigated EEG-based control of a number of devices, including the roll-axis motion of a simple flight simulator [20]. A task display in the simulator (which included a light source flickering at 13.25 Hz) provided a random series of commands requiring the operator to roll right or left to specific target angles. The operator accomplished this control by raising the evoked response above a high threshold to roll right and suppressing the response below a low threshold to roll left. In this example, the control system produced a discrete output when the amplitude of the evoked response remained

above an experimenter-specified threshold for 75% of the samples in a one-half second interval. This combination of threshold and duration criteria required the user to produce sustained changes in the response; however, brief fluctuations did not interrupt system control. A typical simulator control trial is shown in Figure 11. Most subjects were able to acquire 70-85% of the roll-angle targets after 5-6 hours of training (Figure 12).

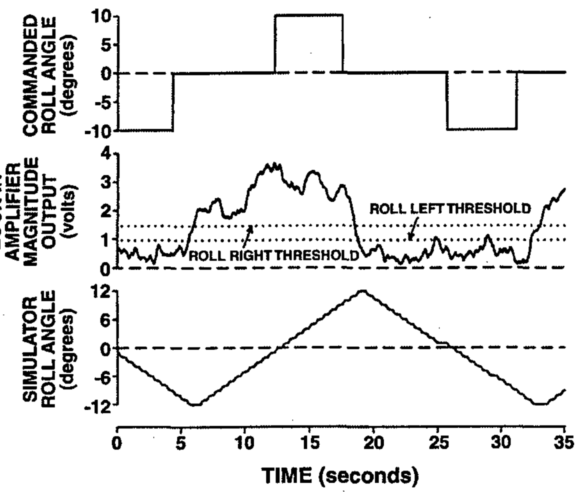


Figure 11. Control of simulator roll-axis motion using the visual evoked response. The lock-in amplifier provides a continuous measure of the magnitude of this response. Responses above one threshold produce motion to the right (positive), while responses below a lower threshold produce motion to the left (negative). From [20].

In addition to the simulator control application, the same group employed evoked response control to operate a neuromuscular stimulator designed to exercise paralysed limbs [21]. Raising the evoked response above a high threshold turned the stimulator on and suppressing the response below a low threshold turned it off. This, in turn, caused the user's knee to extend or flex in response to changes in stimulator current. A series of specific knee extension angle commands was presented in each trial to test user performance. A group of three subjects was able to acquire 96% of the knee angle targets presented in a brief pilot study.

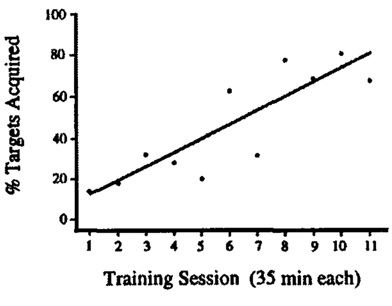


Figure 12. Learning curve for one subject performing the roll-axis motion control task using the visual evoked response. Subject had no prior biofeedback or simulator training. Data points are means of 16 trials in each session. Solid line is a linear regression on these means. From [20].

4.5.1.3. Discussion

The principal difference between the two EEG self-regulation methods discussed above is the presence or absence of an evoking stimulus. In both approaches the user controls the amplitude of a brain signal, but in one case the fundamental signal is evoked by external events. The use of an evoking stimulus complicates the interface design and requires that some of the user's sensory and perceptual resources be devoted to the processing of this input. In addition, the evoking stimulus may serve as a distraction or be poorly accepted by some users. On the other hand, the evoking stimulus produces a time-locked EEG response. This permits one to use synchronous signal processing techniques that improve noise tolerance and reduce or eliminate the confounding effects of other activities and rhythms. An open question concerning the self-regulation of internally-generated brain rhythms is the applicability of this approach with active, multitasked users; how difficult will it be to discriminate intentional and natural variation? At the present time, it is premature to discount either of these approaches based upon such considerations. Only further development and application will identify the real constraints associated with each method.

Both of the self-regulation approaches require significant calibration or adjustment for individual users early in the training process. Once users establish reliable control, the calibration values tend to remain quite stable from day to day.

4.5.2. Processing Natural EEG Responses as Control Signals

Several approaches to EEG-based control are based upon the interpretation of spontaneous brain responses. Using this approach, little or no user training is required. Informal observations suggest, however, that overall system performance may improve with experience, i.e., users may develop the ability to enhance their spontaneous responses in order to improve their control.

4.5.2.1. Evoked Response Level Coding

Natural variation in the amplitude of brain responses evoked by external stimuli can be used for control. One example is the P300 component of the event-related potential (ERP). The P300 is a positive-going component of the ERP response to a sensory input, with an amplitude in the 5-10 microvolt range and a latency of about 300 milliseconds. This response is most prominent over the central and posterior (parietal) regions of the scalp. Many studies have demonstrated that the P300 is enhanced when users are presented with a stimulus that is of low probability or has special significance. If, for example, a user is asked to select a particular item that is presented in a series of items, the user will produce a larger P300 when the desired item is presented [22].

Unfortunately, it is not possible to reliably discriminate among P300 responses to single presentations of a series of items. Multiple presentations and response averaging are required. Farwell and Donchin [23] investigated the rate of stimulus presentation, the number of presentations, and the type of signal processing algorithm while using the P300 response as a means for subjects to select one element from a 36 element matrix of letters and words. In this case the stimuli were repeated intensifications of the rows and columns of the matrix. Farwell and Donchin found that they could discriminate among the P300 responses using interstimulus intervals as short as 125 milliseconds, which caused the responses to overlap. Despite this overlap, a minimum of 26 seconds was required to generate discriminable responses to each of the 36 matrix elements. As a result, a communication rate of 2.3 characters per minute was the best that they were able to achieve.

Another spontaneous response that has been evaluated for EEG-based communication is the visual evoked potential or VEP. This microvolt-level signal is produced by visual stimuli such as flashes and colour reversals. The major components of the transient VEP occur within 80 milliseconds of the stimulus, and are most commonly measured over the posterior (occipital) region of the scalp. As with the P300 response, multiple presentations and response averaging are required to estimate the amplitude of the VEP.

Sutter [24, 25] used the VEP as a means for subjects to select elements from an 8 by 8 matrix of letters and words presented on a computer monitor. The matrix elements were modulated in intensity or colour and the VEP to each modulated element was individually computed. His approach was based on the fact that a modulated stimulus in the centre of visual field evokes a much larger VEP than one in the visual periphery (Figure 13). The system selected the character with the largest response as the desired one, i.e., the one the user was visually fixating.

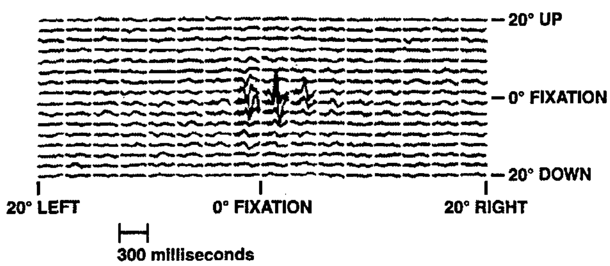


Figure 13. VEP to central and peripheral visual stimuli. From [24].

A powerful methodological aspect of Sutter's approach was the use of m-sequences (white pseudo-random binary sequences) to control the elements of the flickering matrix and to extract the average response to each matrix element from the combined signal. This signal included hundreds of overlapping VEP responses. The use of m-sequences allowed Sutter to generate discriminable responses to each of the 64 elements in approximately 1.5 seconds. Each of these responses was correlated with a reference VEP template collected in a 10-20 minute preliminary session. These correlation coefficients were then compared to each other and to a threshold value. If coefficient n remained above threshold and was larger than all others for a specified amount of time, then matrix element n was selected.

By using m-sequences and virtual keyboard overlays, communication rates of 10-12 words per minute were achieved. For example, the first keyboard overlay contained the alphabet and many frequently used words. If the desired word was not on that screen, the user selected the beginning letter, which brought up an overlay of words beginning with that letter. Clearly, Sutter's approach provided practical communication rates, but potential users must deal with a somewhat unpleasant flickering keyboard display.

4.5.2.2. Pattern Recognition

Rather than focusing on the amplitude of a single EEG response, control can be based on more complex spectral, temporal or spatial patterns in the EEG. For example, virtual joystick operation has been demonstrated by Hiraiwa, Shimohara and Tokunaga based upon neural network recognition of the EEG patterns that precede joystick movement [26]. Following network training (as many as 1000 repetitions of each movement), the authors were able to predict the direction of joystick movements with 96% accuracy. This evaluation was conducted with one subject and off-line analysis of the data. By way of comparison, the authors also attempted real-time prediction of the utterance of one of two Japanese vowels and reported 100% success after 1000 network training trials.

Alternatively, one may focus on more specific patterns of EEG activity associated with the cortical preparation for body movements. One such pattern is the reduction in mu rhythm (8-12 Hz) power in the sensorimotor area of the cerebral hemisphere contralateral to the movement [27, 28]. Pfurtscheller and his colleagues have attempted to use this and other such patterns to classify finger, toe or tongue movements before they actually occur [29, 30]. In particular, they have focused on power decreases, event related desynchronisation (ERD) in the 8-12 Hz band and brief power increases, event related synchronisation (ERS) in 30-40 Hz band (Figure 14).

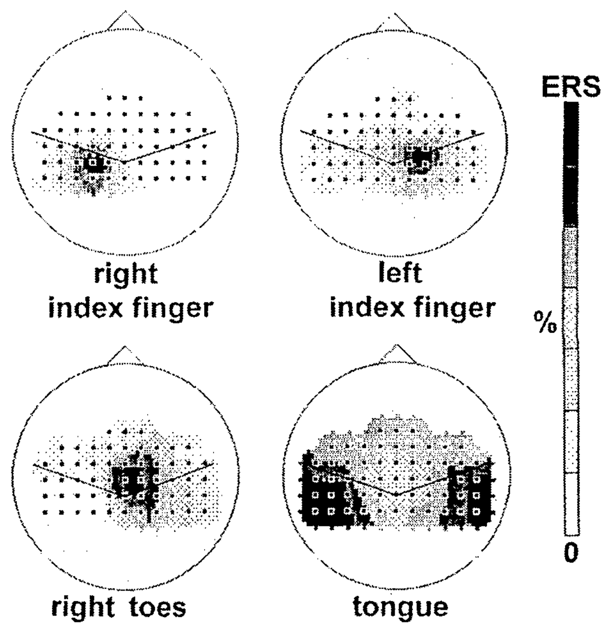


Figure 14. Brain maps showing specific patterns of event related synchronisation (ERS) associated with the preparation for specific body movements. From [30].

These signals were recorded with an array of 8-14 electrodes spread over the central and posterior regions of the scalp and processed using DSP techniques. Classification was accomplished with Kohonen Learning Vector Quantization (LVQ) [31] which iteratively defined a set of reference vectors for each classification category, in this case finger, toe or tongue movement. Following training, these reference vectors were used by the network to classify new EEG input vectors. Pfurtscheller and his colleagues typically used only 100-200 trials to train the LVQ, far less than the number employed by Hiraiwa. The greater temporal and spatial specificity of the EEG patterns being classified by Pfurtscheller may be a major contributor to reduced network training time.

Pfurtscheller's off-line system achieved 89% accuracy in predicting button pushes with the left or right hand. With toe

and tongue movement added, accuracy dropped to 70%. In addition, the neural networks could be trained with imagined rather than actual movements, but movement prediction was slightly degraded in this case.

While Hiraiwa and Pfurtscheller provided their neural networks with samples of EEG collected at successive time points, their static neural networks did not actually assimilate the temporal dimension of the patterns being classified. With static networks, the successive samples are provided as simultaneous inputs to separate nodes of the input layer. Barreto, Taberner and Vicente [32], have recently begun to evaluate the potential of dynamic neural networks for the classification of EEG patterns that represent the preparation for body movements. With dynamic neural networks, the input consists of a temporal sequence of values provided to a single input node. Such networks store past samples of the inputs in memory structures that perform a time-to-space mapping for the classifier.

4.5.2.3. Discussion

The evoked response approaches require no training of the user, or of the signal processing algorithms. Since the responses are essentially time-locked to an external stimulus, selecting temporal windows for signal processing is readily accomplished. However, the requirement to average multiple responses, in order to obtain reliable amplitude estimates, can be a significant constraint. Sequential presentation of the evoking stimuli limited Farwell and Donchin [23] to very low character selection rates. Sutter's [25] use of m-sequences permitted highly-overlapping stimulus presentation and significantly improved the output bandwidth of his system.

The pattern recognition techniques, which all employ neural networks, require individual training of the recognizer. As noted above, the neural network must be trained with 100-1000 repetitions of the EEG patterns to be classified. One potential approach to this training issue is to conduct it implicitly rather than explicitly. For example, the user might continue to physically operate the controls, while the EEG-based recognizer observes the brain patterns associated with these activities. Once the recognizer can satisfactorily predict certain control actions, it could then be permitted to take over those functions. This implicit training might be conducted in simulated or synthetic environments, for example.

Practical application of the pattern recognition techniques must address the issue of selecting the temporal window that includes the EEG patterns to be recognised. This problem is analogous to the challenge faced by speech recognizers operating in a high noise environment. The experiments conducted to date create an artificial solution to this problem. The user is given explicit cues to execute the movements to be predicted from the EEG, and the pattern recognizer is synchronised to these cues. In the real world, such cues typically will not exist. Barreto et al. [32] argue that dynamic neural nets will reduce this problem, but this advantage has not been demonstrated in a real world environment.

Finally, individual differences represent an additional constraint on both approaches. Evoked response amplitudes, dynamic EEG patterns, artefact characteristics and optimal electrode locations all vary from person to person. The pattern recognition techniques tend to address these issues during recognizer training, while the evoked response approaches often require initial tuning of electrode locations, signal processing algorithms and response templates for each individual. Fortunately, many of these sources of variation are fairly stable from day to day, once the signal processing parameters have been optimised for each user.

One can also compare the approaches based on EEG self-regulation with those that employ spontaneous EEG responses. The former are clearly less natural and intuitive than the latter, since the user must produce artificial changes in their EEG. This does not mean that such changes are inappropriate, interfere with other cognitive activities, or are difficult to produce. Rather, users must learn to produce these changes, and this requires an investment in training. Once these EEG patterns are under voluntary control, there is a great deal of flexibility in how these patterns can be applied. Essentially, their application is constrained only by the bandwidth, resolution and accuracy of EEG self-regulation.

4.6. USER FEEDBACK REQUIREMENTS WITH EEG-BASED CONTROL

Biofeedback is one of the key technologies that enabled the development of systems based on learned control of EEG responses. User feedback has been implemented in two ways: (1) as an inherent part of the task, e.g., movement of the display element being controlled by EEG, or (2) as a separate display element when movement of the controlled element does not provide timely information. Although biofeedback is not required in systems that are based on the recognition of naturally occurring EEG patterns, it is still possible that such feedback will allow users to improve the speed and accuracy of their EEG responses.

5. FUTURE DEVELOPMENTS

This lecture has discussed a wide variety of tasks that employ biopotential signals for control. With creative interface design, almost any discrete response task can be performed with these modalities. In certain cases, response time advantages have been demonstrated using EMG signals to replace physical movement of a conventional control. The ability to perform continuous proportional control is clearly much more limited. Difficulty in producing and maintaining graded EMG and EEG signal outputs is the reason for this limitation. To achieve this type of control, investigators most commonly employ time-proportional techniques in which the position or velocity of the controlled element is proportional to the amount of time that a biopotential signal remains above an established threshold.

A clear limitation of the current state of the art, with the exception of prosthetic device control, is that little work has been done outside the laboratory. There is a profound need to identify applications that require hands-free operation and to develop biopotential controllers for field evaluation. Only then can developers determine how to achieve effective performance in real-world, multitask, multicontrol environments.

Improvements in signal acquisition hardware are required to support such applications. Dry electrode systems that do not require skin preparation or electrode creams are essential. These electrodes must work on hairy skin or scalp areas and be tolerant of slippage or movement. Self-administered user calibration approaches, as well as signal monitoring schemes that continually adjust the interface based on signal quality and background noise, are required.

Signal processing enhancements are also needed in many areas. While discrete EMG and EEG responses can be identified rapidly, recognition of complex patterns is more time consuming and constrains the speed of human-system interaction. Nevertheless, the pattern recognition approaches offer the greatest potential for discriminating signal from artefact and for controlling multiple DOF systems. In the area of EEG-based control, for example, it is apparent that complex

pattern recognition will be the method of choice if we are to develop true "intent-based" interfaces.

Finally, as discussed by McMillan, Eggleston and Anderson [33], biopotential control may be most applicable when used in a manner that bridges the space between operator state monitoring and explicit control. Referred to as an intelligent controller paradigm, this approach employs an intelligent interpreter that monitors a range of human outputs, including EMG and EEG signals, to infer user intent, a desire for information, and so on. The interpreter then issues commands to the system, consistent with user intentions. Recognition of the EEG patterns that precede specific physical movements is a simple example of this notion. Detection of a specific EEG response permits the interpreter to infer that the user desires to push a specific button. The intelligent interpreter approach, rather than simple substitution of EMG and EEG signals for conventional inputs, represents the path to achieve optimal utilisation of biopotential-based control.

6. ACKNOWLEDGEMENTS

Much of the material on EMG-based control was prepared by Dr. Bernard Hudgins, University of New Brunswick, Canada, as part of the final report of AGARD Working Group 25, entitled Alternative Control Technologies. The author wishes to thank Dr. Hudgins for his kind permission to use this material.

7. REFERENCES

1. Parker, P.A., and Scott, R.N., "Myoelectric Control of Prosthesis", CRC Critical Reviews in Biomedical Engineering, 13, Issue 4, 1986, pp 283-310.

2. Scott, R.N., and Parker, P.A., "Myoelectric Prostheses: State of the Art", Journal of Medical Engineering and Technology, 12, 1988, pp 143-151.

3. Hogan, N., and Mann, R.W., "Myoeletric Signal Processing: Optimal Estimation Applied to Electromyography - Part 1: Derivation of the Optimal Myoprocessor", IEEE Transactions on Biomedical Engineering, BME-27, 7, 1980, pp 282-295.

4. Scott, R.N., Paciga, J., and Parker, P., "Operator Error in Multistate Myoelectric Control Systems", Medical and Biological Engineering, 16, 1978, pp 296-301.

5. Paciga, J., Richard, P., and Scott, R.N., "Error Rate in Five-State Myoelectric Control Systems", Medical and Biological Engineering, 18, 1980, pp 287-290.

6. Hudgins, B., Parker, P., and Scott, R.N., "A New Approach to Multifunction Myoelectric Control", IEEE Transactions on Biomedical Engineering, BME-40, 4, 1993, pp 82-94.

7. Patterson, P.E., and Katz, J.A., "Design and Evaluation of a Sensory Feedback System that Provides Grasping Pressure in a Myoelectric Hand", Journal of Rehabilitation Research and Development, 29, 1, 1992, pp 1-8.

8. Scott, R.N., "Feedback in Myoelectric Prostheses", Clinical Orthopaedics and Related Research, 256, 1990, pp 58-63.

9. Sullivan, G., Martell, C., Weltman, G., and Pierce, D., "Myoelectric Servo Control", Report to US Air Force Aeronautical Systems Division under Contract #AF33(657)-7771, May, 1963.

10. Clark, J.E., and Phillips, S.J., "The Efficacy of Using Human Myoelectric Signals to Control the Limbs of Robots in Space", NASA-CR-182901, 1988.

11. Farry, K.A., Walker, I.D., and Baraniuk, R.G., "Myoelectric Teleoperation of a Complex Robotic Hand", IEEE Transactions on Robotics and Automation, 12, 5, 1996, pp 775-778.

12. Fernandez, J., Farry, K.A., and Cheatham, J.B., "Waveform Recognition Using Genetic Programming: The Myoelectric Signal Recognition Problem", Genetic Programming 1996 Conference, 1996.

13. Junker, A., Berg, C., Schneider, P., and McMillan, G.R., "Evaluation of the CyberLink Interface as an Alternative Human Operator Controller", US Air Force Technical Report AL/CF-TR-1995-0011, 1995.

14. Nelson, W.T., Hettinger, L.J., Cunningham, J.A., Roe, M.M., Lu, L.G., Haas, M.W., Dennis, L.B., Pick, H.L., Junker, A., and Berg, C.B., "Brain-Body Actuated Control: Assessment of an Alternative Control Technology for Virtual Environments", in "Proceedings of the 1996 Image Conference", 1996, pp 225-232.

15. Vodovnik, L. "An Electromagnetic Brake Activated by Eyebrow Muscles", Electronics Engineering, 1967, pp 694-695.

16. Vatikiotis-Bateson, E., Munhall, K.G., Kasahara, Y., Garcia F., and Yehia H., "Characterizing Audiovisual Information During Speech", in "Conference on Spoken Language Processing - CDROM", 1996, Paper No. 1010.

17. Kang, W., Cheng, C., Lai, J., Shiu, J., and Kuo, T., "A Comparative Analysis of Various EMG Pattern Recognition Methods", Medical Engineering Physics, 8, 5, 1996, pp 390-395.

18. Wolpaw, J.R., and McFarland, D J., "Multichannel EEG-Based Brain-Computer Communication", Electroencephalography and Clinical Neurophysiology, 90, 1994, pp 444-449.

19. Wolpaw, J.R., McFarland, D.J., Neat, G.W., and Forneris, C.A., "An EEG-Based Brain-Computer Interface for Cursor Control", Electroencephalography and Clinical Neurophysiology, 78, 1991, pp 252-259.

20. McMillan, G.R., Calhoun, G.L., Middendorf, M.S., Schnurer, J.H., Ingle, D.F., and Nasman, V.T., "Direct Brain Interface Utilizing Self-Regulation of the Steady-State Visual Evoked Response", in "Proceedings of the RESNA 18th Annual Conference", 1995, pp 693-695.

21. Calhoun, G.L., McMillan, G.R., Morton, P.E., Middendorf, M.S., Schnurer, J.H., Ingle, D.F., Glaser, R.M., and Figoni, S.F., "Functional Electrical Stimulator Control with a Direct Brain Interface", in "Proceedings of the RESNA 18th Annual Conference", 1995, pp 696-698.

22. Donchin, E., Karis, D., Bashore, T.R., Coles, M.G.M., and Gratton, G., "Cognitive Psychophysiology and Human Information Processing., in Coles, M.G.H., Donchin, E., and Porges, S.W., (Eds), "Psychophysiology: Systems, Processes, and Applications", New York, Guilford Press, 1986.

23. Farwell, L.A., and Donchin, E., "Talking Off the Top of Your Head: Toward a Mental Prosthesis Utilizing Event-Related Brain Potentials", Electroencephalography and Clinical Neurophysiology, 70, 1988, pp 510-523.

24. Sutter, E.E., "The Visual Evoked Response as a Communication Channel", in "Proceedings: IEEE Symposium on Biosensors", IEEE, 1984, pp 95-100.

25. Sutter, E.E., "The Brain Response Interface: Communication through Visually-Induced Electrical Brain Responses", Journal of Microcomputer Applications, 15, 1992, pp 31-45.

26. Hiraiwa, A., Shimohara, K., and Tokunaga, Y., "EEG Topography Recognition by Neural Networks", IEEE Engineering in Medicine and Biology, 19, 1990, pp 39-42.

27. Jasper, H.H., and Penfield, W., "Electrocorticograms in Man: Effect of the Voluntary Movement upon the Electrical Activity of the Precentral Gyrus", Arch. Psychiat. Z. Neurol., 183, 1949, pp 163-174.

28. Chatrian, G.E., Petersen, M.C., and Lazarete, J.A., "The Blocking of the Rolandic Wicket Rhythm and Some Central Changes Related to Movement", Electroencephalography and Clinical Neurophysiology, 11, 1959, pp 497-510.

29. Pfurtscheller, G., Flotzinger, D., Mohl, W., and Peltoranta, M., "Prediction of the Side of Hand Movements from Single-Trial Multi-Channel EEG Data using Neural Networks", Electroencephalography and Clinical Neurophysiology, 82, 1992, pp 313-315.

30. Pfurtscheller, G., Flotzinger, D., and Neuper, C., "Differentiation between Finger, Toe and Tongue Movement in Man Based on 40 Hz EEG", Electroencephalography and Clinical Neurophysiology, 90, 1994, pp 456-460.

31. Kohonen, T., "The Self-Organizing Map", Proceedings of the IEEE, 78, 1990, pp 1464-1480.

32. Barreto, A.B., Taberner, A.M., and Vicente, L.M., "Classification of Spatio-Temporal EEG Readiness Potentials towards the Development of a Brain-Computer Interface", in "Proceedings of the 1996 IEEE SouthEastcon Conference", IEEE, 1996, pp 100-103.

33. McMillan, G.R., Eggleston, R.G., and Anderson, T.R., "Nonconventional Controls", in Salvendy, G., (Ed) "Handbook of Human Factors and Ergonomics, 2nd Edition", New York, NY, Wiley, 1997, pp 729-771.

Human Factors Issues for the Integration of Alternative Control Technologies

Dr G.M. Rood
Systems Integration Department, Air Systems Sector
Defence Evaluation and Research Agency, Farnborough
Hampshire GU14 0LX, UK

SUMMARY

The introduction of Alternative Control Technologies (ACTs), and their closer links with human natural behaviour, will require a better balance between the human factors requirements and the aircraft integration engineering issues. Successful integration of ACTs into aircraft systems should provide significant operational advantages, and the following paragraphs discuss an approach for the necessary balance of human factors and engineering.

1. INTRODUCTION

Current control in cockpits generally uses only manual switching and this has been a traditional method from the earliest days of aircraft use. Technology has advanced to such an extent that there now are a number of alternative ways of entering or inputting data and information into an aircraft system. However, like most technology insertion into real aircraft systems, if it is not integrated properly then significant problems can occur in service use. The use of these alternative control technologies utilises the more natural ways of human communication and requires that a human centred approach to integration be used to a greater extent than is currently used in design.

Thus there becomes two main approaches that must be reviewed in the integration of Alternative Control Technologies (ACTs).

* Human factors approach

This examines the ideal design process and discusses the human factors tools that are available to develop, refine, and evaluate interfaces. This approach is designed to capitalise on the strengths and minimise the weaknesses of human operators.

* An engineering framework

This examines mechanical and electrical issues associated with the selection and location of new components in the crew station: and the computational architecture required to interpret nonconventional control outputs from the human and integrate them with the outputs from conventional controls.

Both segments to this approach are considered complementary, can be pursed in a concurrent fashion and have, in fact, been used, to a greater or lesser extent, in a number of US, CA & UK helicopter programmes (LHX, Apache, Comanche, EH101/Merlin, Kiowa etc.)

2. HUMAN FACTORS APPROACH

The design process does not, unfortunately, only involve the more rational aspects, such as engineering and technology, but includes the more imponderable aspects like commercial, political, pragmatic, managerial and end-user pressures. In the areas of human interface design, operator or pilot requirements should retain a major influence. ACT may be used for reasons other than human factors issues, such as workspace limitations, but there remain many human factors methods that will help integrate the new control technologies

Preferably, however, the decision to use ACTs within an interface should be made as the best operator-centred solution within the existing engineering constraints.

Design processes are becoming more multi-disciplinary and concurrent, and many design drivers, such as useability and maintainability, can influence a design at far earlier stages than was previously possible. Ideally a human-centred design philosophy should be adopted to ensure that a system is designed starting with the operators interface and using human factors principles. Although final design will always be a compromise, the best human factors practices should be used and an understanding of how much system performance is lost by restricting the human engineering aspects.

As human-in-the-loop systems become more complex, the humans' task can also become more complex, requiring more interaction with the system, and bottlenecks may occur at the interface. ACTs may help in these 'overload' conditions and allow an increase in the effective communication bandwidth between the system and human, thus allowing a greater amount and complexity of information exchange.

2.1 ACTs as Supplements and Substitutes

ACTs may be used, in many cases, as supplements and substitutes within an existing interface. Supplements are used when, for example, a new task is being introduced, or it may be a substitute for another control. For instance, voice control can be used to switch radios instead of manual selection, but in both cases of supplements and substitutes it is important to analyse the whole set of concurrent and subsequent tasks which the operator has to carry out and not just focus on the local context. Thus it is almost essential to use some form of task analysis to ensure that all of the inter-relationships have been accounted for.

The introduction of a form of ACT into an existing system is considerably simpler than with a hypothetical design, since it involves factors that can be observed and measured and compared in terms of effectiveness or performance against the existing system. But the choice and implementation should be made with the understanding of the demands that

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

the task place upon the human operator, and the alternative ways in which it is possible for the operator to do the task.

2.2 Future Interface Developments

It is difficult in current practices in many complex systems, military or otherwise, to achieve a man-machine interface that is based on human characteristics alone. This is certainly so for current systems, but, in future systems, where technological development allows - such as in Virtual Environment Interfaces (Crewstations etc) - the limitations of physical factors, such as space, cost & safety, are liable to be less restrictive. The interface is only one aspect of the human machine integration and the relative roles of the human and the machine are evolving. The machine is now increasingly capable of sophisticated behaviour, and therefore the contexts and rules of the interaction could become quite complex. Currently most human operated systems are just that - a master/slave relationship, where command inputs have a fixed structure and meaning. It is feasible that the operator and machine could work more like a team, such as supervisor and operator, where the machine is capable of inference, adaption to changing circumstances and complex decision making. In these type of cases the human-plus-machine would become, in effect, a joint cognitive system. The potential is for powerful and sophisticated performance, but each system may have incomplete information about each others intentions, similar to human-to-human team work. The interface is thus more dependent on the communication of uncertain and potentially ambiguous information. Cognitive ergonomics has arisen as a human factors discipline to address specifically the integration of the users psychology and the systems information processing, and the hope is that progress in this field will help resolve some of these significant problems. ACTs have an important role in making such interfaces possible by broadening the scope of human-machine communication. But it is the interpretation of an ACT output that, in many cases, requires both considerable human engineering knowledge and engineering development.

2.3 The Benefits of Human Factors and Human Centred Integration.

One of the primary benefits of bringing human factors into the design process is the necessity to methodically analyse the human-machine interface, and that forces an analysis of factors that are too easily taken for granted or overlooked. With a man-in-the-loop system there are a number of permutations and unknown factors which can affect the way in which tasks are carried out, and the human engineering process uses various models and techniques which can ensure that certain factors are addressed, and can represent and predict some aspects of human performance in the system. These models and techniques have to deal with imprecise, incomplete and uncertain data, yet provide a useful input to the design process. This is accomplished by combining the knowledge of human physiology and psychology with experimental studies and using methods which force designs to make allowance for the range of variance which is probable for a particular population of operators and the particular conditions in which they will carry out the task. This is less critical in the design of current systems, where controls are usually manually switched and, in general, the information required may be only the ability to reach the control and the time taken to complete the control task. The use of ACTs will, however, require a greater knowledge and understanding of physiology and psychology, and, in particular, data will be required about natural behaviour (eye and head-movements, body movements, non-verbal communication etc), response times and sensitivities in different modalities, and sensory-cognitive compatibility. This will allow the ability to prescribe optimum design solutions rather than analyse and assess pre-specified options. These user-centred designs could offer radical solutions by identifying unconventional interface techniques, and, if these are not technically possible, will provide a useful aim for technology development - a possible balance of the technology push-pull process.

2.4 Human Factors in the Design Process

In the current world much of the human factors design for aerospace related interfaces is for modifications to existing designs. Whilst this human engineering process is important to go through, the advantages are liable to be effective but marginal.

If a new design is being offered, the comprehensive interface design process could involve the following steps;

* Identify top-level task requirements (e.g. Mission Analysis).
* Analyse and model the task (allocation of function i.e what the operator will do and when; what the machine will do and when -- but not how).
* Determine what communication needs to take place between human and machine.
* Develop recommendations for interaction requirements (dialogue) based on the type of communication and context.
* Develop initial recommendations for interface technology, based on type of interaction requirements.
* Develop initial design specifications for the interface content based on detailed task analysis, human factors guidelines and predictive models of human performance.
* Produce a (rapid) prototype of the design.
* Evaluate the design with user trials to establish how the operator does perform his allocated functions.
* Re-iterate as required to achieve the required human performance.

The last four steps should be a part of any interface design process, and as many of the preceding steps as possible included. There are a number of human factors tools used in the interface design process and these may include: task analysis and modelling methods, human performance models and databases, guidelines, design philosophies, simulation, experimental investigation and human performance metrics.

2.5 Review of Human Factors/Engineering Tools

2.5.1 Design principles and frameworks

In human engineering design there are human factors 'principles' which have been compiled to provide high level aims when designing an interface. Schneidermann (1), for instance, describes eight main principles:

* Dialogues should be consistent.
* Systems should allow short cuts through some parts of familiar dialogue.
* Dialogues should offer informative feedback.
* Sequences of dialogues should be organised into logical groups.

* Systems should offer simple error handling.
* Systems should allow actions to be reversed.
* Systems should allow experienced users to feel that they are in control, rather than the system is in control.
* Systems should aim to reduce short term memory load (users should not be expected to remember much).

Similarly Dix et al (2) describes three principles which can be summarised as:

* **Learnability** - take advantage of natural behaviour to reduce learning needs.
* **Flexibility** - use general (flexible) interaction rules rather than task specific.
* **Robustness** - systems should provide transparent feedback so that errors are understood.

Other guidelines principles include those of Williges et al (3), who offer seven dimensions (Compatibility, Consistency, Memory, Structure, Feedback, Workload and Individualisation) and a substantial (679) set of guidelines from Smith and Mosier (4), parts of which are relevant to integrating ACTs. A breakdown of the dialogue parameters provides a framework for analysing the sort of interaction which is required by the task and allows the designer to determine whether a particular control method can be implemented to match a particular type of task.

Nominally there are five aspects to any interaction;

* Style: how fixed the interactions are (i.e. interfaces may be constant or adaptive).
* Structure: how constrained the rules are (e.g. protocols or natural language)
* Content: how explicit or implicit conveyed information can be (e.g. semantic codes or raw data).
* Context: how context dependent the interaction is, including contexts such as system failure).
* Mode: traditionally only two modes have been identified: verbal and spatial. This may be insufficient for interfaces with ACTs, where implicit pilot state modes may be used.

McMillan, Egglestone and Anderson (5) describe two types of paradigm for coupling operator intentions to machine activation;

* The Servo Paradigm: effectively a monologue rather than a dialogue, which involves the operator making pre-determined, intentional commands to invoke fixed machine responses.
* The Structural Coupling Paradigm (SCP): this views the operator as a performer, whose performance is monitored in order to ascertain what the machine should be doing.

In the use of ACTs, implementation may be under either paradigm, but the SCP will require ACTs that can monitor the performer. Also, the SCP requires other sources of information, for instance about the vehicle status and an inference engine to interpret the appropriate action to instigate the machine. In this, more complex, framework, the operator becomes just another variable which has to be taken into account by the system to determine the most appropriate action.

The choice of coupling paradigm will depend upon the particular circumstances into which an ACT is being integrated. The simplest is the servo paradigm, and probably more appropriate for 'upgrading' systems where servo coupling is already in use. The SCP is more appropriate as an overall framework for the whole of the control interface.

2.6 Allocation of Function

The obvious allocation of function is as to capability, such as human workload capacity or decision making capability, data processing ability etc. But, if this is the only consideration, the overall performance of the human and system may be sub-optimal. But there are other considerations such as maintaining alertness, job satisfaction, retention of training, error minimisation or avoidance and crew interaction which may have important implications for allocation of function.

When ACTs are bought to the interface, the allocation of function should refer not only to the sharing of tasks between the human and machine, but also to allocation of tasks to different sensory modalities.

There are a number of approaches to allocation of function and at least four ways in which functions are allocated:

* allocation to machine by 'a priori' management decisions.
* allocation according to respective capabilities.
* allocation by formal analysis of tasks and sub-tasks.
* allocation by Fitts' list.

General good practice precludes the allocation of as many tasks as possible to the machine, as this risks leaving the human out of touch with the machine. The first two methods noted above are appropriate for tasks where there are constraints that cannot be removed (ie only the operator can make the decision to attack a target). The use of Fitts List involves looking up a specific function (eg data sensing) and reading off a list of pros and cons for human and machine performing that function. However, this method generally precludes the ability to show how a single task can be shared between man and machine, and cannot be used for the more advanced approaches to interface design which may be generally described as 'joint cognitive systems'.

As in many of the human engineering analyses, there is no magic formula to prescribe optimum allocation of function and also these techniques do not handle dynamic or adaptive allocation.

In many cases, allocation of function can only be carried out by comparing a number of potential solutions, by the use of operator trials and/or the use of modelling aspects of human performance/workload, errors, etc.

2.7 Task Analysis & Modelling

The creation of a formal and auditable trail in the early parts of the design process can be accomplished through a formal representation of the tasks, which will, in turn, allow interdependencies to be assessed, problem areas to be identified, and important aspects of the task to be taken into account. The representation can be based upon an analysis of different factors, such as processes, functions, goals, human knowledge or skills.

Taken to an extreme task analysis can create large amounts of unwieldy data and decomposition paths should be taken only as far as is cost effective. STANAG 3994 and US MIL-H-46855B specify a 'critical task analysis' to be carried out on those tasks which are predicted to have high workload, or which are critical to safety or mission success.

2.8 Taxonomies

Taxonomies delineate the categories or classes into which a task or activity can be separated, such as actions, skills, performance, knowledge etc. They are important as they

allow an achievement of consistency and repeatability in the analyses, especially with respect to level of detail.

For ACTs, an appropriate taxonomy would define the sorts of function each ACT could usefully perform (eg track, select, indicate stress etc). It would also be useful if the taxonomy could also capture the features by which ACTs could benefit interaction eg 'track target (eyes), showing that the eyes always looked at the object to be tracked: this highlights a possible exploitation of natural behaviour, such as eye pointing. There is thus a need to develop a control taxonomy especially for ACTs.

2.9 Human Factors databases

The use of general human factors databases must be undertaken with care. Much of the information is shown in the form of experimental results, many of simple dual mode interactions, and the application to more specific complex integration issues can be difficult. But they provide an invaluable source of information, particularly for the physiological (eg perception) thresholds and limits, and provide good guidance, but may not be directly applicable to a given interface problem.

The major Human Factors database available can be found on CASHE PVS, a CD-ROM tool including the Engineering Data Compendium (Boff & Lincoln (6)), MIL-STD-1472D and a Perception and Performance Prototyper.

2.10 Predictive Modelling

There are many human performance models, created for different purposes and with varying degrees of validation. Some aspects of human performance, particularly those in the sensory modes (Visual, Auditory) are more amenable to validated modelling than others (eg cognitive). They provide a useful first pass at the earlier stages of interface design. AGARD WG22 is developing an expert system, HOPE (Human Operator Performance Evaluator), to assist in the selection of Human Performance models.

There are many tools, basically very similar, for predicting workload with different task and interface designs. Most are based on Wickens (7) multiple resource theory and generally involve a task analysis being performed, either on the basis of a design proposal or from observation of an existing task, and an assessment of workload over time is generated. This allows an identification of potential overload or underload problems. Such tools include POP (Predictor of Operator Performance from DERA, UK; W/INDEX (Workload Index) from Honeywell, USA; PUMA (Performance and Useability Modelling Tool) from Roke Manor Research, Siemens, UK; WINCREW from Micro Analysis and Design, USA.

There are also models of anatomy and biomechanics such as SAMMIE and 'Jack' and are implemented in a computer based environment into which geometric information about the workstation can be imported.

Further tools attempt to provide a more integrated tool for human factors analysis, but are inevitably more complex and require more expertise to both run and make use of their outputs. These tools are difficult to comprehensively validate due to their relative complexity, but provide a useful design tool if used correctly. A number of those type of predictive modelling tools is noted below and are described in more detail in NATO Technical Report AC/243: A Directory of Human Performance Models and System Design.

* HOS (Human Operator Simulator).
* EPIC (Executive Process-Interactive Control).
* COGNET (Cognition as a network of Tasks).
* IPME (Integrated Performance Modelling Environment).
* MIDAS (Man-Machine Integrated Design and Analysis System).

2.11 Error Modelling

The assessment of the likelihood of error is important for any interface, but there will be particular considerations for the use of ACTs. ACTs are essentially designed to make use of 'natural' human behaviour and this behaviour is perhaps more prone than other interface behaviour to contextural influences, and intrinsically more variable. This is one of the reasons why redundancy is more important, ensuring more than one behaviour can be used to control an input. Variability of behaviour can occur between individuals, but also within an individual over time. In this context is will be important to understand both error occurrence and error recovery. In many cases it may be more efficient to design a system which allows rapid and efficient error recovery, rather than try and reduce error probability to an acceptably low value. The benefits to operational performance may ameliorate the potential problems of producing higher error rates, if error correction is appropriate, rapid and efficient.

Conventional methods of assessing error rates (error analysis by activity analysis, subject matter experts, previous data etc.) will not necessarily be applicable to ACTs, and it will be important to take an approach to error assessment which firstly can identify the potential cause of errors and the context dependency of those causes, and secondly allows task based probabilities of errors to be mapped onto the causation model.

2.12 Rapid prototyping

It is preferable to have completed the analysis and design of the interface before starting rapid prototyping, otherwise the assessment can only be on a hit-and-miss basis. If limited options exist, rapid prototyping can be used to compare those designs, and to assess human performance. But the value of the exercise is dependent upon how the evaluations are made. Rapid Prototyping makes use of tools to assist in creating an adequate representation of the interface for evaluation purposes, and these tools include VAPS (Virtual Prototypes Inc.), Designers Workbench (Coryphaeus Inc), or Virtual Reality, or some simple physical mock-ups with behind-the-scenes human substitutes for machine functionality.

2.13 Evaluation and Performance Measures

Evaluation of each integrated ACT is different, and evaluation trials must be tailored to particular requirements, conditions of use etc. It is important to identify, at the outset, what criteria are important for the assessment and how they can be measured. Some criteria for ACTs may include:

* **Compatibility**: perhaps indirectly measured by ease of learning, error rates, intuitiveness, reduced workload, better situational awareness.
* **Capability**: faster response times, greater accuracy, capacity for parallel activities.
* **Reliability**: Fewer errors, less variability.

* **Flexibility**: Ease of reconfiguration, and reallocation, versatility achieved by operators.
* **Acceptability**: User ratings, trials in workplace, analysis of socio-cultural context.

Some of the measures, such as timing and errors (speed and accuracy), are easily measured, but any measures can be meaningless unless they are made within the context of careful experimental design which takes into account the user sample, the control of variables, the order in which the tests are made, the way in which experimental participants are briefed and the statistical techniques to evaluate the results.

The measurement of workload or situational awareness would be very valuable in assessing the impact of the whole interface: the sum of its various components. Several tools are being developed to try and do this.

* **NASA TLX** (Task Load Index) and variants, is a NASA developed tool in the form of a paper or computer based questionnaire, to assess subjective impression of workload.
* **SWAT** (Subjective Workload Assessment Technique), developed by Wright Patterson Air Force Base, is an on-line subjective assessment tool with 3 'domains' of workload, each of which is given a rating between 1 and 3 at critical parts of the task. Recent developments with this technique have enabled the identification of a 'red-line' maximum workload, above which performance drops off.
* **SAGAT** (Situational Awareness Global Assessment) was developed by Northrop as a direct measure of situation awareness, but as it requires task interruptions, it is of limited use, and may be unacceptable to users in evaluation trials.
* **SART** (Situational Awareness Rating Technique), from DERA, is a questionnaire for assessing subjective situation awareness, which has been refined through repeated use. There are several versions, including CC-SART which aims to assess the cognitive compatibility of interfaces.

There are others measures worth considering, particularly in the physiological domain. Eye-movement patterns, blink rate, heart rates, Galvanic skin response, EEGs (both a.c. and d.c.) and other aspects of secondary task performance, training time, behaviour modification - all can provide some measure of performance. However, it is important to assess the variance of such measures both within individuals over time, and between different individuals.

It must be emphasised that the design and evaluation process should be regarded as an iterative cycle which starts by putting together a human-machine interface which has be recommended through an analysis of requirement. It involves conducting a series of evaluations, primarily so that the design team can understand its useage and eliminate unforeseen defects before the arrangement is frozen and built. The foregoing material discusses key ergonomic and psychological issues and describes a variety of tools that can materially assist this process. Ultimately the design team must use their judgement in choosing which issues to address and in selecting the most appropriate tools for their application.

3. ENGINEERING INTEGRATION

The satisfactory introduction of novel controls into an working aeroplane introduces a gamut of engineering issues, and it is assumed that the procurement of the equipment would be carried out to comply with the wide range of engineering standards applied by the customers, often, in military aircraft case to MIL standards in the USA, Defence Standards in the UK, or, increasingly, to commercial standards. The introduction of ACT will, however, necessitate some reconsideration of the physical arrangement of the cockpit systems and the flow of information between the cockpit and the remainder of the aircraft systems.

3.1 Mechanical and Electrical design

The schematic, Fig 3.1, shows how the integration of novel controls could involve mounting of components in the airframe, the cockpit, on aircrew clothing and on the pilot. Most aircraft-mounted equipment is rack mounted and generally high density packed, and retro-fits become a problem, but not always insoluble. Future systems will use modular avionics, which should allow additional facilities to be accommodated relatively easily, and such architectures are intrinsically reconfigurable and to a high degree fault tolerant.

The use of an electro-magnetic helmet and hand tracking system, to provide head movement tracking and gesture or virtual pointing control, the transmitter would be the only component that needed to be cockpit mounted. This would probably be bonded to the inside of the canopy just above the rear of the helmet. Optical tracking would, however, require more sensors and different locations to minimise reflections and avoid the capture of direct sunlight.

A major consideration in the fitting of systems to operational and experimental aircraft is the question of emergency escape by the use of the ejection seat. The elements that are attached directly to the crew member or his helmet and clothing, must be arranged to separate automatically from the other cockpit mounted or airframe mounted units,

as well as allowing the further separation of the pilot from the seat later in the ejection sequence. The additional problems of more electrical cabling and connectors will become more problematical and alternative approaches, for instance the use of multiplexed fibre optic channels to transfer data in digital form, become increasingly attractive.

The design of aviators helmets, incorporating both night vision devices and sensor displays in addition to the original protective, life support and communication functions is already a serious challenge, primarily because this integration must be accomplished without increasing headborne mass. The incorporation of head and eye tracking systems, and later biopotential sensing systems on the helmet will need careful design and consideration to minimise any mass increases or increasing CofG offsets.

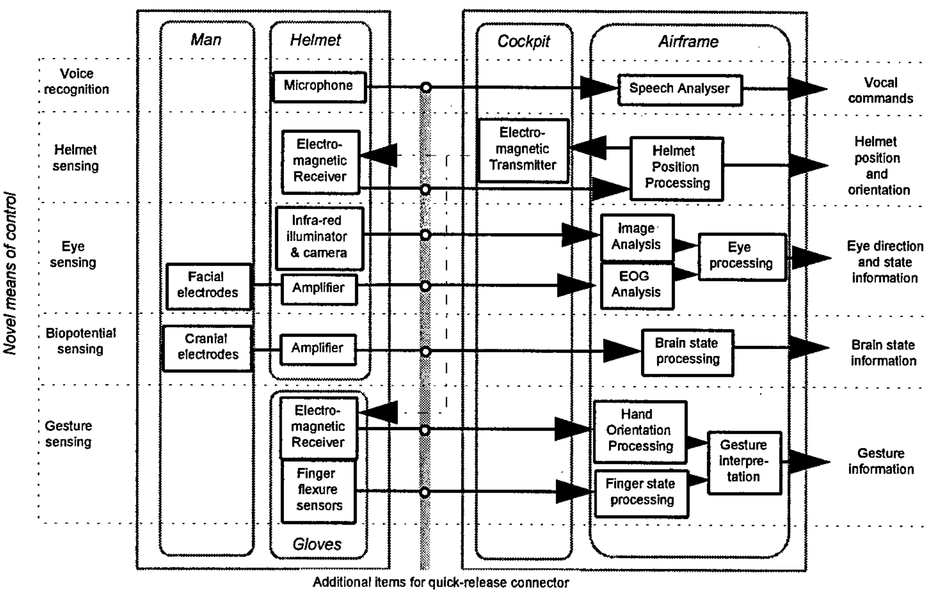


Fig. 3.1 A schematic representation of the location of novel control system components

3.2 Computational

Having arranged the satisfactory physical installation of the novel control suite, it is necessary to connect it to the rest of the aircraft avionics. Future aircraft systems will inevitably include some advisory aids, 'intelligent' or otherwise, and some pilot state monitoring, in addition to comprehensive mission, utility, flight control and weapon systems. One approach would be to integrate the novel control suite with the conventional systems, shown schematically in Fig 3.2.

The additional function, called the 'command interpreter', adjudicates between the signals generated by any of the novel or conventional control modalities in order to send an unambiguous command to the relevant aircraft system. Such a command interpretation function is already incorporated in advanced fighters, for instance Eurofighter and Rafale, primarily as a means of integrating the voice control system in these aircraft.

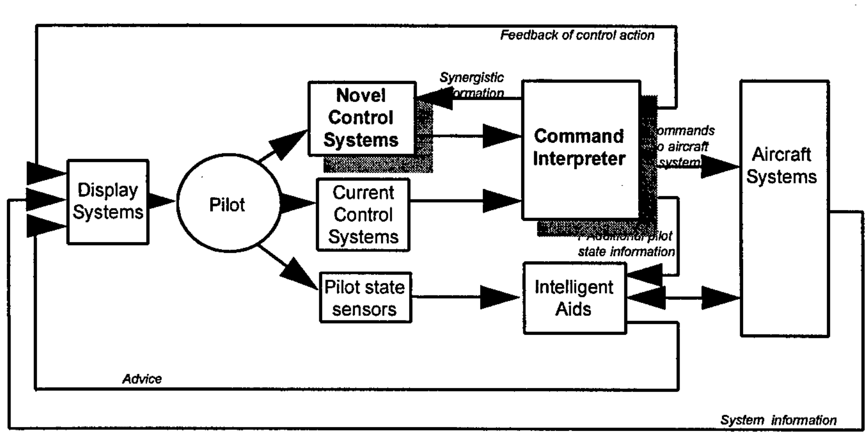


Fig. 3.2 Broad information flow paths in an aircraft with novel control systems

The fundamental requirement is that the command interpreter should accept only the intended inputs: switch selections, utterances, designations, gestures and perhaps physiological and mental state measurements of the pilot. Intention, as accepted by the interpreter, being defined by the pilot, by doctrine, by tactics and by other factors. Unintended utterances, designations, gestures and mental and physiological states should be identified and have no effect on the aircraft systems. This could be accomplished by software in which the allowable links between the received output from the controls and the signals which are sent to the systems could be specified as a set of finite state 'rules'. These would trap errors, such as double selections, and express the constraints and flexibilities which, by analogy with man-computer interaction, constitute the aircraft operating system. For instance, the ability to select an external object by fixating with the eye, pointing the helmet sight or indicating by a hand pointing gesture, then saying 'target', 'lock radar'or 'range', or the flexibility to perform a mixture of these actions, would be programmed as a set of command acceptance rules.

The command interpreter could produce three classes of output in addition to the main collated system control signals. These would be:

1) Information fed back to the displays system so that the user can be kept aware of the state of the control mechanisms in order to operate them satisfactorily. For instance, this would include the highlighting of a virtual key selected by eye fixation, the visual and auditory presentation of the output from speech recognition systems, and the movement of a cursor responding to finger pointing.

2) Synergistic feedback to the novel control suite, to enhance the performance of one system using information derived from another sensing system. For instance, speech recognition reliability could benefit from knowledge of eye pointing direction by biasing the context of the objects or selections near the pilots eye fixation.

3) The novel controls produce additional pilot state information, for instance the pilots eye pupil diameter and his blink rate from the eye tracker, his formant frequency from the speech recogniser and his head activity from the helmet tracker. All of these would supply extra information which could assist the intelligent aid which monitors the pilots state to make a more reliable classification, for instance whether workload has induced boredom or frenzy, and whether he was conscious.

3.3 Control of the Controls

Conventional control mechanism all have fixed characteristics and cannot be matched to the qualities of the operator - indeed the human factors specifications set out, to a large extent, the 'standard' human (and the range) such that the individual qualities of the humans are minimised. In contrast, ACT will probably need 'control controls' so that they can be set up to suit the individual user and be de-selected in the event of a failure.

The most evident need is to be able to calibrate the appropriate sensing system to match the possible mixture of voice, eye, hand gesture and cortical response characteristics of the user to optimise accuracy and reliability. Any time consuming calibrations need to carried out on the ground, perhaps in a simulator, and any pre-flight checks should be simple, quick checks to confirm correct system function and set an alignment, and it should not be necessary to engineer facilities which allow re-calibration in flight.

Finally, the high level means of exercising control over the ACTs could be engineered by something as simple and unambiguous as a dedicated panel housing a short row of 'on/off' toggle switches. If the user becomes convinced that, for instance, spoken commands are being interpreted erroneously, he can switch the voice recognition system off, knowing that the command interpreter will be aware of this de-activation, and he can continue the mission using the remaining facilities. This would be true of all the ACTs as they will, at least in the near future always be ***alternative*** ways of controlling the aircraft systems and not necessarily the primary method, unless chosen by the pilot to be so.

4. REFERENCES

These references are only a selection of those 34 references in the WG 25 paper proper.

1. Schneiderman, B. 'Designing the User Interface: Strategies for Effective Human Computer Interaction', Masachusetts, Addison Wesley 1987.

2. Dix A, Finlay J, Aboud G, and Beale R 'Human Computer Interaction', Hemel Hempstead, UK Prentice Hall 1993.

3. Williges R C, Williges B, and Elkerton J. 'Software Interface Design', in Salvendy G (Ed), 'Handbook of Human Factors', New York, NY, Wiley 1984.

4. Smith S L and Mosier J N, 'Design guidelines for user-system interface software', The Mitre Corporation Technical report ESD-TR-84-358, 1984.

5. McMillan G R, Eggleston R G, and Anderson TR, 'Non-conventional Controls' in Salvendy G (Ed) Handbook of Human Factors and Ergonomics, 2nd Edition, New York, NY, Wiley 1997.

6. Boff K R and Lincoln J E, 'Engineering Data Compendium, Human Perception and Performance', Vols 1,2 &3.

7. Armstrong Aerospace Medical research laboratory, Ohio 1988.

8. Wickens C D, 'The structure of processing resources' in Nickerson R and Pew R (Eds) ' Attention and Performance VIII, Hillsdale, NJ,Erlbaum.

Synthesis-and expected benefits Analysis

Alain Leger

SEXTANT Avionique/MMI

BP 91 - Rue Toussaint Catros

33166 Saint-Médard-en-Jalles CEDEX - France

SUMMARY

A synthetic approach of the various Alternative Control Technologies is proposed, taking into account advantages and inconveniences for military aircraft applications. Operational rationale, classification of technologies following capabilities and degree of maturity, summary of main functional characteristics and integration issues are critically reviewed. A brief presentation of multimodal dialog issues is also presented. Finally, a tentative investigation of potential areas of benefits for military aircraft design and operation is conducted.

1. INTRODUCTION

From a purely theoretical standpoint, allowing the human operator to expand control possibilities on aircraft systems beyond simple manual actions constitutes definitely a significant advance in term of Man-Machine communication improvement. Implementation of effective non-conventional « Alternative Control Technologies », such as voice or gaze control, would *ideally* let the operator use intuitively his own communication strategies, then generating considerable benefits. Moreover, such benefits could apply to both side of the man machine interface.

On the human side, using intuitive communication strategies rather than arbitary mechanical actions may allow to minimize the « cost » of interaction with the system, in selecting the most situationally adapted control modality(ies). This way, sensorimotor, attentional and cognitive costs could be optimized , function of operators' intentions and external constraints of the moment. Globally, better use would be made of the limited resources of human beings. A positive effect could also be found on training needs, but this remains to be demonstrated.

On the machine side, generalization of virtual controls would result, at least in new aircraft, in a drastic reduction of the number of dedicated mechanical switches and control panels. The resultant gain in space would then facilitate cockpit layouts including very large size displays. The interactivity provided by the new controls is expected to play an essential role in regard of the usability of such large displays. The alternative control devices are highly susceptible to be integrated in modular avionics systems and existing equipments as Helmet Mounted Displays. The replacement of bulky and quite expensive control panels by these new controls may have some positive impact on cost, especially if applications are developping in the public domain. Maintenance costs could also be substantially reduced. It is quite common that mechanical switches are jammed or even broken by pilots strictly abiding to the old principle « If it jams, force it; If it brakes, anyway it needed replacement ». Controllers as speech recognizers should be less susceptible to that kind of problems and more easily serviceable than mechanical control panels.

From a practical standpoint, enthusiasm for these new technologies has to be quite tempered and the state of the art review shows that things are not so simple.

In the first place, experience and also some surveys tells us that pilots usually take quite conservative positions when asked about using new control devices. Such a reserved position is easily understandable, since manual control has been exclusively used from the beginning of aviation, in most case satisfactorily. Manual control is robust, extremely reliable and the large variety of controllers developed by engineers covers quite adequately the various classes of usage found in combat aircraft. There is also a strong consensus among pilots that physical contact with the control device generates a high level of confidence.

To overcome the users' legitimate concerns, the need to introduce new technologies, alternatives to manual controls, should be carefully analyzed and

Paper presented at the RTO Lecture Series on "Alternative Control Technologies: Human Factors Issues", held in Brétigny, France, 7-8 October 1998, and in Ohio, USA, 14-15 October 1998, and published in RTO EN-3.

clearly demonstrated in regard of the tasks and activities to be performed. Collective expertise and common MMI « know how » is far to be enough in this domain and a strict methodological approach is probably the only way to avoid further disenchantment.

Above that, it has to be recognized that most available technologies suffer from more or less stringent technical limitations, sometimes hampering seriously their usability or, at least, increasing the cost and difficulties of operation through complex installation and functional procedures. In most case, a significant gap also exist between the human capabilities (in term of range, speed, accuracy, semantic content of the signal,....) and the sensing technique used to capture these capabilities (movement, speech, biopotentials). As an example, the most technically mature techniques in speech recognition (probably also among most mature Alternative Control Technology) appear still very crude in regard of the richness of natural human language.

So far, operational considerations and applications, technological and integration issues have been individually reviewed. Our purpose will be now an attempt to summarize, synthesize and compare the main advantages and limitations of the various technologies, with regard to the various control issues onboard aircraft. The potential areas of benefits which could be expected for crew station design will also be addressed critically.

2. SYNTHESIS ON OPERATIONAL AND TECHNOLOGICAL ISSUES

Introducing non-conventional technologies as an alternative to conventional manual control automatically brings several questions: Do we really need to do that? What is the level of maturity of the various technologies, their advantages and limitations? Does the integration of these new technologies requires specific attention? should these technologies rather be used stand-alone or in a cooperative (multimodal) way?

Answering these questions is of course intimately linked to the intended application. The data previously presented in the different lectures aims to provide support to designers regarding this matter. Some general issues deserve, however, to be pointed out.

2.1 The need for alternative control technologies

Most « new » control technologies are in fact around for quite a while now. Generations of engineers and scientist have worked on gaze tracking systems, speech recognizers and other non-conventional controls. Besides some of these technologies have reached a certain degree of technical maturity, there appears to be currently some emerging operational motivations justifying more general usage of non-conventional controllers.

Very demanding operations in the current generations of fixed and rotary wing aircraft considerably increase the need for « eyes out » operation particularly at night and in poor weather. Meanwhile the complexity of aircraft systems and the speed of operation require from the pilot to constantly interact with the system in order to properly configure it and acquire information. Twenty years ago, pilots were still flying manually the aircraft and faced with the need to rapidly access information, without spending too much time head down and while keeping their hands on the main controls (Stick and Throttle). The HOTAS concept, introduced in the seventies, brought for a while an acceptable solution to this problem. Despite massive introduction of flight control automation, current trends show that we are clearly approaching the limits of this concept, particularly in regard of the overload of pilot's short term memory. Excessive number of HOTAS accessible functions induces difficulties, especially in highly time constrained situations. Such problems are susceptible to result in increased error rates in issuing commands and additional requirement for training. Anthropometric aspects have also to be considered. The larger number of switches required to access the HOTAS functions clearly raises a space problem, interfering with the design of the main controls. Difficulties in handling correctly all the switches are already reported by some « small » pilots and this will definitely be aggravated with the arrival of female crew in the cockpit.

On the positive side, non-conventional controls may also offer new possibilities, which could not be obtained without great difficulties neither by mechanical manual controls, nor by aircraft mounted sensors. The best example, if not the most useful, is probably the capacity to fire missiles with large off-boresight angle given by Helmet Mounted Sights. The head tracking technology used in these helmet has been widely demonstrated to accurately cue missile seekers on target located 60° off-boresight or beyond. Using gaze tracking technology, without the need to display a reticule, would make more natural, easier and faster such designation, particularly under G-loads or in time-constrained situations. It would also considerably expand the aiming envelope. The implicit use of these technologies in head-slaved sensors systems, such as those currently used for night vision in combat helicopter, constitute probably the best example of current operational use of alternative control. Similarly, the ability of voice control to enter a complex hierarchical control structure at any point constitutes a feature which cannot be easily matched using conventional controls.

The operational rationale to introduce Alternative Controls Technologies onboard aircraft appears

twofold. It would constitute a way to alleviate current problems. and offer solutions for the future, as it is expected that more complex systems will almost inevitably require more control mechanisms. It would allow the pilots to perform more efficiently using unique features offered by non-conventional controls.

2.2 Technological issues

Obviously, there is very large differences in the degree of maturity of the reviewed alternative control technologies, ranging from operationally fielded (head trackers in helicopters and fixed-wings) to pure research laboratory (EEG). The control capabilities of the various technologies shows differences, from discrete Cursor Control Devices (CCD) to high level communication. Before summarizing advantages and limitations of these technologies, it appears therefore of interest to introduce a simple classification in regard of what they are good at and degree of maturity.

2.2.1 Classification

Basically, all alternative control technologies have a CCD capability. two classes could be introduce here, discrete and continuous. Only some technologies, Direct Voice Input (DVI), gesture and brain control (EEG) are high level communication capable. Table 1 shows an attempt to classify these technologies following their capabilities, with reference to their maturity level as they are already flying, R&D mature or still in research laboratories.

Function / Technology	Cursor Control Device Capability		Communication Capability		Technical maturity
	DISCRETE	CONTINUOUS	CURRENT	POTENTIAL	
Touch Screen	●	-	-	-	High
Touch Pad	●	●	-	-	High
Head-Tracker	-	●	-	-	High
Eye-Tracker	-	●	-	-	Medium
DVI	●	-	●	*	High
Gesture	●	●	●	*	Low
EMG	●	●	-	?	Medium to low
EEG	●	●	-	●	Low

Table 1: Capability and maturity of alternative control technologies (●: exist, -: non-existing, *: potential improvement)

Touch screens and touch pads are contact devices. They are border between Alternative Control Technologies and manual control. With variations due to the realization technologies, Touch Screens are preferably used to enter discrete inputs, when Touch Pads are used as a mouse or a joystick. Head and eye trackers are considered here as continuous CCD as they require an additional validation input (mechanical switch or else) to perform designation. Duration of fixation on an object would be difficult to use for validation, since there is no on/off position for head and eye signals. DVI can be considered to have a discrete capacity as a pointing device but tracking is unpracticable. It has the capacity of high level communication and still a large potential for improvement. Gesture has both discrete and continuous CCD capability, associated with communication capability through signs language. It definitely appears as the most complete input channel, unfortunately technical maturity remains low and limitations aboard aircraft are quite severe. EMG has only CCD capabilities and has been shown to allow both discrete on/off inputs and continuous tracking. Through sophisticated signal processing EEG has the same CCD capabilities, but still lacks of real communication capability. Should complex pattern recognition software be developed for EEG control, this technology would then potentially offer the basis for true « thought-based » interfaces. Maturity of this technology is of course low.

2.2.2 Summary of advantages and limitations

The different technologies will be reviewed sucessively in order to summarize and comment the main characteristics of each one, essentially in regard of military cockpit applications.

2.2.2.1 Touch screens - Touch pads

Touch pads and Touch screen are typical Cursor Control Device designed to operate in « Glass Cockpit ». Touch Pad positioning accuracy has been shown to be worse than input devices such as trackball, but they are definitely faster to operate. On this last aspect, Touch Screens are far better than Touch Pads but their accuracy is considerably worse. Comparatively, Touch Screens are also considered as more comfortable, intuitive and procuring the least fatigue to operate, but Touch Pads remain quite acceptable in regard of these subjective criterion. Touch screens inputs are known to be more affected by turbulence than Touch pads.

Flight test results in the « Rafale » have shown that the touch pad associated with a collimated Head Level Display (HLD) has been very rapidly an intuitively used by pilots in all flight conditions. Location, positioning accuracy and adequation of the size of the touch pad was found satisfying. Due to technical difficulties, the Touch Screen lateral LCD displays required more time to be usefully evaluated. Once these difficulties were adequately solved, the level of satisfaction of pilots was good and they start using routinely the touch screen to activate the menus on the displays. It has to be noted that visual and haptic feed-back was available in the latest prototypes versions.

There is an obvious complementarity between the two devices. When possible, it may be interesting to consider an heterogeneous redundancy of such devices. That mean that the complementarity of devices such as Touch Screens and Touch Pads could be used to optimize pilot actions. Pilots would be then free to use the best input modality in regard of own preferences, task to be performed and environmental conditions

2.2.2.2 Head and Eye trackers

When used explicitly, Head and Eye Tracker are basically of the CCD type. Main functional difference with Touch Screens, Touch Pads and other mechanical pointing devices is that they are not attached to a specific display area. They can virtually access all locations in the surrounding environment.

Currently, head trackers are used in Helmet Mounted Sight and Displays (HMSD) for direct and reverse cueing (operator/system or system/operator) of target direction. The HMSD Line of Sight (HLOS) could also be used alternatively to a contact device to control a cursor on cockpit display, provided adequate parallax correction is made. This function would be quite interesting with very large display, where it may be difficult to visually locate a manually controlled cursor. However, accuracy of head trackers is not very good in regard of tasks requirements and it is difficult to keep the head stationary, particularly in dynamic environments. HLOS can be used implicitly to avoid clutter when the pilot looks head down in the cockpit, as a control input to blank unnecessary HMD imagery or symbology. Implicit control of head-steered sensors, especially for night vision, has been successfully demonstrated. In this case, accuracy and bandwidth of most trackers are largely sufficient, which is not always the case for the dynamics of the sensor platform. It has recently been shown in helicopters during NOE flight that pilots' head peak velocity could reach 240 °/s, widely exceeding most current sensor platform performance. Most currently mature head tracking devices use Electro-Magnetic of Electro-Optical techniques, providing a reasonably good accuracy and dynamic characteristics. Improvements should be brought to these techniques in regard of robustness to environment perturbations (respectively sunlight and metal parts). A significant improvement in static and dynamic accuracy is required to allow head mounted virtual cockpit application, in particular virtual HUD.
It has already been pointed out by others that, by many aspects, pointing with the head is quite non-natural. Actually, eye and head movements are strictly physiologically coupled during everyday life activities and Gaze (resultant of eye + head vectors) has clearly been shown as the controlled variable for the Central Nervous System. Using a head pointing device deprives the operator from the benefit of this physiological coupling, in term of angular coverage, speed, accuracy and stability of visual fixation. Naturally stabilized by vestibulo-ocular mechanisms, the gaze line of sight, usually referred as Point of Gaze (POG), is also less likely to be affected by turbulence and sustained accelerations in combat. As POG is the controlled variable, it would not be necessary to display an eye-slaved cursor or reticule to designate a point in space. Continuous secondary visual feed-back (presentation of own point of gaze as measured by the gaze tracking device) has been described as more disturbing than helpful in some situations. Theoretically, using POG rather than HLOS in controlling cursor or designating target should present many advantages, as it would be enough to « look » at object and validate to complete the intended action. Unfortunately, the accuracy of « usable » eye trackers (Corneal reflection/pupil) is not very good (~ 1°).That means that all task requiring a great accuracy, as selecting a way point on a navigation display could not be completed using eye or gaze tracking alone. It has to be noted that, for a given location in space, the errors of eye and head trackers are not combined linearly in the resultant POG accuracy. Actually, eye/head coordination mechanisms are such that the combination of eye and head movement to reach a given point allow the respective trackers to operate in better conditions than for eye or head alone. Other eye-tracking techniques may have better accuracy, however, they are usually totally unacceptable outside the laboratory environment. It has to be recognized that all available eye trackers are quite difficult to operate, even in laboratory conditions. At the moment, this technology is only mature in the R&D domain and in benign environments as flight simulators, provided skilled personnel supply assistance for the necessary settings. Significant progress in term of robustness of the measurement process, opto-mechanical integration in head gear and automatization of adjustments and calibration procedures are required before gaze tracking becomes flightworthy in combat aircraft. Work is underway in several countries in this domain and current available technology in optics, sensor and processing should allow to achieve the necessary enhancements.

2.2.2.3 Direct Voice Input

Direct Voice Input is commonly presented as the most mature Alternative Control Technology. Indeed, a large amount of work, including numerous flight tests, have been devoted to develop the different components of this technology. Still, there is no system operationally fielded, which in regard of this criteria makes head tracker technology the only really mature. The 25 years of development of Voice control system in aeronautics have been marked by successive waves of enthusiastic optimism usually followed by pessimistic periods. We are in a high now, as progress achieved during these last years allow to be reasonably optimistic about the effective implementation of this technology on several programs (EFA2000, Rafale and JSF). DVI would then become the first non-

conventional control with high level communication capability to be implemented on combat aircraft. It has to be noted that the use made of speech control remains limited in regard of its natural richness, as only semantic content of speech signal is used in current recognizers. Other kinds of information, as emotional effects or cues to control the dialog with another speaker are considered as perturbations, though may be of interest to enhance detection of pilot's intention.

Currently, continuous speech, speaker-dependent systems are quite readily available for military aerospace applications. Vocabulary size about 200 words and branching factors 6/8 (syntax perplexity) are quite commonly considered as suitable for fighter aircraft. Of course, the nature of the intended application, in terms of the characteristics of users, tasks to be performed and environment, plays a primordial role in determining the most adapted combination of vocabulary size and syntax perplexity to obtained the required performance.

Automatic Speech Recognition techniques (ASR) essential functional elements are signal acquisition, signal processing and pattern matching. The two last components have reached now a quite good maturity level, even if some progress potential is existing. Some attention is currently focused on the signal acquisition step, before starting signal processing. This point appears currently as a real challenge in current speaker-dependent systems and will be even greater in speaker independent applications. Assessing the performance of ASR systems is usually based on speech recognition rate determination in various conditions, using different speakers. These rates are expressed in terms of word or sentence recognition rate (respectively, WRR and SRR). SRR is preferably used in military cockpit applications, since it better reflects the robustness of the ASR system in regard of the whole system performance. Currently SRR rates are, for recent studies, in the range of 90 to 97/98 %, especially if several utterances are considered in case of error on the first one. On the human factor side, one point to be mentioned is the dependence of speaker's performance on habituation to the device and environment. Such effects were found quite clearly apparent, both in laboratory studies (including centrifuge trials) and flight tests.

Besides working on the essential components of ASR, several ways can help to improve voice control. As it is quite likely that a small error rate will still persist despite technical improvements, it seems important to provide to the user good feed-back on system recognized command and allow easy correction on error detection. Offering both auditory and visual feedback to the pilot appears quit adapted to aerospace operations. Specific correction commands as « delete », « correction » or « insert » are usually provided to the users, but more elaborated solutions are possible through dialog modeling. Additional sources of information as automatic leap-reading systems have been shown to further enhance the robustness of speech recognizers and some work is underway in this area. Numerous environmental factors are susceptible to affect speech production, rendering ASR more difficult. In the aerospace environment, the effects of ambient noise, physical (G-loads, vibrations) and emotional stressors have been quite extensively studied during these last years and are now better understood.

Some improvements are highly desirable in term of global robustness of DVI systems but other problems should also be addressed. Though using speech is supposed to be easy and intuitive, the way speech recognizers currently works if far to be optimal in regard of natural speech usage. Speech is naturally a quite slow communication process. Pilots express quite clearly this concern by stating that, when the situation starts really to get tense and dense, it becomes very difficult to organize speech in a rigid way, until the moment speech is even too slow to follow the action. Speech recognizers are themselves quite slow in processing the commands and add to this problem. This introduce a serious limitation to the use of DVI in muddled and time constrained situations, when a fast and intuitive communication channel would be most useful. The need to follow a rigid syntax and remember a precise vocabulary is also hampering the use of voice control. Introduction of « speech understanding » systems capable of understanding any command, however it is phrased, would considerably improve this aspect. Last but not the least, additional attention should be paid both to speaker-independent systems and use of DVI in a multi-speaker environment.

2.2.2.4 Gesture

Gesture appears definitely as a very powerful control channel, both in regard of CCD capability and high level communication potential. Gesture can provide discrete static inputs as well as generates dynamic complex commands. It is a very intuitive and natural communication capability in humans, with no doubt widely anterior to the acquisition of articulated speech. With exception of some head and body movements, the semiotic function of gesture is mostly concentrated in movements of hands and upper limbs. Actually, gesture can be seen as a total « virtual » equivalence of classical manual control (except for the haptic part), with in addition considerable high level communication capabilities. Numerous scientific teams have well understood the importance of gesture as a communication tool in highly computerized environments. Scientific publications on this topic are usually of excellent quality, showing the interest elicited by gesture-based control.

Despite the remarkable possibilities of gesture, limitations of this technology are quite severe, particularly in regard of military aerospace applications. Among these limitations are fatigue, effects of G-loads and vibrations, lack of feed-back, repeatability and variability of gesture. Besides, as

head and eye movements, it is difficult to identify start and stop of a gesture, though homogeneous validation is possible (without requiring another input channel). From a practical standpoint, the cohabitation of gesture with the HOTAS concept seems extremely difficult, if not totally incompatible. Using a gesture-based control in a combat aircraft would probably imply that every control would have to become virtual, including HOTAS controls, which is clearly unacceptable with current state of the art technology.

Above these difficulties, the main inconvenience of gesture based control is the low maturity of technical solutions used to capture gesture. Most devices are intrusive and interfere with users freedom of movement. Non-intrusive techniques as video cameras, magnetic or optical devices still need considerable improvement before providing required range, accuracy and reliability. Gesture based control could probably be used with some benefits in more benign environments than aircraft cockpit as control rooms or operational centers.

2.2.2.5 Biopotentials

Use of biopotentials represents a quite fascinating area in the Alternative Control Technology domain. EMG and EEG signal processing to provide control input currently elicits a considerable interest in advanced research laboratories, specially in the US. EMG signals have been used for quite a bit of time for prosthetic device operation and this kind of technique is recognized now to have a significant clinical value. EEG-based signals are currently under investigation. It has to be observed that most current work in this domain mainly relies on a very clever use of signal processing software, rather than explicit neurophysiological considerations. Still mostly in the phenomenology domain, results looks surprisingly promising as it appears quite easy to control various dynamic process or even fly a simulator with such techniques and minimal training.

Current EMG and EEG-based control systems are clearly limited to CCD functions. They basically carry a potential for communication as both EMG and EEG signals might be used for early detection of pilots' intents. Though many applications could be suggested, including the very mediatic « fly by mind », the practical use of such devices in a cockpit seems quite remote. Severe technical limitations exist currently both for the capture of the signals, requiring intrusive contact electrodes, and in regard of current signal processing capabilities. Should these limitations be overcome and breakthrough on complex pattern recognition of EEG signal really be achieved, then the fast and intuitive communication channel required by some pilots would become available. For now, true « thought-based » interfaces are probably far closer to dream than reality. It has, however, to be borne in mind that such technology would probably supercede all the others if available one day.

2.3 Integration issues

A reasonably good corpus of knowledge exists on most current usable technologies. Despite this knowledge, individually accumulated on each technology, it as to be considered that the determination of the « good practice » to integrate these technologies is still in its infancy. So far, only CCD like devices have been integrated, as HMSD head-trackers on combat aircraft or touch pads and touch screens. Most of the integration efforts have been spent on mechanical and electronic system integration, human engineering considerations remaining, so far, more implicit than explicit in the integration process. Difficulties can be expected when attempting to integrate communication capable technology in increasingly complex aircraft systems.

System integration issues can be split up between two axes: human factors and system engineering considerations. Without sufficient attention to and coordination between these two domain domains, there is little chance that a successful integration could be achieved and potential benefits of Alternative Controls fully delivered. Even an experienced design team with good knowledge of operational tasks and conditions may have considerable difficulties to achieve the integration process, within an acceptable level of industrial risk, without to have recourse to sound human factors methodologies. Both in regard of human factors and system engineering , two cases should be considered, use of non-conventional controls as supplement or substitutes in an existing cockpit and totally new interface development.

In the first case, attention should be paid to the tasks that the operator has to carry out, not only those affected by the new control mode, but also in regard of possible indirect effects of this new control on the whole system operation. The analysis of already existing tasks is, however, relatively easier than predicting entirely new activities. Task analysis should be thorough enough to apprehend all inter-relationships between tasks and the demand that the task places upon the human operator. In the case of a new interface design, where non-conventional controls would be introduced, to obtain operator's performance enhancements, completion of tasks unsuitable to manual control or control simplification, A « human-centered » design process should be conducted. Main steps of such an process should be as follows: identification of top level task requirement, analysis and task modeling, determine the man-machine communications needs, develop recommendations for interaction requirements, develop initial requirements for interface technology, rapid prototyping, evaluation and iteration to obtainment of the required performance.

More than conventional controls, Alternative Control Technologies require this « human centered design » to fulfill the ultimate goal of designing a

true « joint cognitive system ». Along with physiological and psychological knowledge, cognitive ergonomics can help to address the integration problems. The variability which characterize the human being, the impact of imprecise and uncertain data relative to the field of Alternative controls deserve an approach making good use of methods and tools developed by human factors scientists. Evaluation and performance measures should also be carefully conducted following appropriate human factor guidance.

It is quite likely that current design guideline available to system engineers will not be sufficient to cover all the integration issues. Subsequently to the introduction of novel controls, the physical arrangement of the cockpit and the flow of information between the cockpit and the remainder of the aircraft systems would have to be quite substantially reconsidered. This is expected to require innovative approaches and ingenuity from system design engineers. Mechanical and electronical integration of new control devices should be greatly facilitated by introduction of modular avionics systems. Meanwhile, fitting new boxes in the equipment bays of in-service aircraft will remain as usual a challenge. For many of these technologies , a key point to be considered is the validity of the various trade-off design in regard of safety and operational requirements. On the computational integration design, once physical integration as been completed satisfactorily, several points deserve specific attention. The « command interpreter, receiving the input of the conventional and non-conventional controls should have the critical capability to differentiate « intended » inputs from « unintended ». This should be quite easy for DVI when a push-to-talk switch is used, but may be more difficult with some control modes as gesture. Unintended input must be identified and have no effect on aircraft systems. Use of non-conventional controls is not likely to be suitable to operate critical system functions. Software safety issues should, however, be carefully scrutinize and must be treated as crucial to the safety of the vehicle. Feed-back outputs from the systems should also be considered. Informative (to the display system) and Synergistic (in relation with other control modes could be used. Outputs relative to information on pilot's physiological or mental variables could also be used by intelligent aids capable to monitor pilot's state. Last, control of the controls may be necessary to customize the controls settings to match the user characteristics and eventually allow him to set some preferences. It seems also highly desirable to offer to the user the capability to individually control the on/off status of the various alternative controls implemented in the system.

2.4 Multimodal dialog

So far, the various components of what constitutes alternative control technologies have been considered as issuing, sometimes complex, homogeneous command strings when interacting with the system. That means the whole command and is arguments is transmitted using the same control mode and device (monomodal control). Sometimes in everyday life we use cooperation between different control modalities to have a complex action completed by an « intelligent agent ». The most classical example is probably the quite famous « Put that there », where voice and gesture are combined, initially proposed by the Massachusetts Institute of Technology Man-Machine communication research team in the early 80s. Multimodal dialog is therefore defined as the cooperative use of different control modes to interact with a machine. In the case of « Put that there », it has to be noted that a complete command set with arguments could be issued by voice only. However, Voice is known to be quite slow and transmitting in parallel the arguments of the command « put » with gesture may be faster and easier than describing the object, his current and desired location. Somewhere, it has to be considered that the cost of the interaction, function of the local context, is lower using multimodality. That may not always be true, particularly in presence of dynamic perturbations, where the postural control is heavily solicited.

Before continuing it seems convenient to address some terminology issues relative to multimodal dialog. On the operator side, *mode* refers to a psycho-physiological classification, when *modality* is the expression or perceptual orientation used by the operator. As an example, the *modality* « Speech » uses two *modes*, vocal and gesture (lips movements). *Interaction* (man-machine) is the context related use of a specific machine by the operator. *Operator's logic* is the ensemble of natural behaviors used by the operator during a specific interaction. So, a *multimodal interaction* should be broadly defined as an interaction allowing the user to operate the machine following is own logic and not system imposed logic. *Multimodal interaction* is more commonly defined as allowing the operator to combine several modalities to communicate with the machine. On the machine side, *interaction engine* is the logical component of the MMI system centralizing acquisition, interpretation and feed back of interactions between the operator and the machine. *Input media* are the physical devices acted upon by the operator during an interaction. A *multimedia system* should be defined as a system whose architecture allows to manage several input (and output) media.

Of course, many other terms have been defined for multimodal dialog, but the few ones reported above are particularly useful to better understand the « whys and therefores » of this concept. Several classes of cooperation between media are usually described: *redundancy* (same command on different media), *Complementarity* (complementary components of a command on different media), *specialization* (same modality systematically used for a specific command input), *equivalence*

(different media can be selected for an identical input, following operator's preference and context).

If there is no doubt that multimodal behaviors are naturally quite common in normal life situations, the question is to know if this remains true when interacting with a machine. For some authors, following experiments based on the « Wizard of Oz » principle, the complementarity behavior is barely used intuitively. Natural multimodal dialog with a machine would be then more a monomodal, multi-user form of dialog, respecting the preferences and logic of the operators. For others, complementarity remains essential in multimodal interaction. Some interactions using complementarity, as selecting with gaze an object in a large and cluttered display and acting upon it with voice may have a considerable interest when flight management control issues are considered. A any rate, to assess the efficacy of the various classes of multimodal dialog and guide multimedia system design, it may be useful to build metrics reflecting « interaction cost » in terms of sensorimotor, attentional and cognitive demands.

3. AREAS OF BENEFITS

As stated in the introduction, identifying areas of benefits for introduction of Alternative Control Technologies in aircraft cockpit can be seen as obvious or quite controversial following the adopted point of view, theoretical or practical. Basically, there is very little return of expertise on this domain, since these technologies have been, so far, scarcely used in full scale introduction. For the few example available, performance or new control possibilities were sought rather than global task optimization.

Head tracker technology, a key issue for HMSDs, constitutes one of these example. The implicit head control of steerable sensors platform in combat helicopter has been shown clearly now to constitute a real advance for the conduct of night missions. Benefits obtained in tems of operational domain extension are now under evaluation with advanced binocular systems. Conversely, despite many research and flight tests, explicit head control in fixed wing combat aircraft meets more difficulties to find its role in the very sophisticated weapon systems of modern fighters. One of the very few example of an attempt to simplify some cockpit functions is flying in the Rafale, with touch-screen technology on the lateral displays. Technical difficulties were quite high and, though good results have finally be obtained, robustness of such a technology has to be confirmed in the long run. Use of touch screen has allowed to suppress all mechanical function switches usually found around displays, while providing to pilots a very intuitive way to call the different pages on the display. Provided the technology delivers the expected results, benefits in maintenance should also be found on such displays, although paid through additional care requirements for mechanics.

Other technologies may appear now quite rapidly onboard aircraft, as voice control. It is very likely that, in a first time, these technologies will be introduced to supplement manual controls and bring alternative solutions, with little chance to induce cockpit layout simplification. Benefits, however, can be expected in terms of functional simplification of pilot's tasks. Cockpit simplification should be obtained later on, when technologies will be really applied to new interface design.

3.1 redundancy and alternative solutions

The most simple approach, redundancy to already existing manual control, would create a total or partial *equivalence* between the new control (DVI for instance) and the traditional manual modality. The review of integration issues has already stressed the necessary care to be brought on human factors and systems engineering considerations. Potential advantages would be to offer to the pilots an alternative way out, especially when short-term memory problems are encountered with HOTAS switches. DVI, with its capability to access control structure at any point could also bring additional, advantages, but remains slow and could not probably be used in all the HOTAS functions domain. This kind of considerations shows that it is of interest to assess benefits and weakness of the potential technological candidates in regard of human factors and system engineering criteria. Table 2 and 3 show an example, far to be exhaustive, of such assessment for the reviewed technologies, relatively to *current* state of the art characteristics.

Table 2 considers 5 system engineering criteria, response rapidity, as the time between an input to the control system and an output to the system. (fast = 20 ms or less), reliability (makes a consistent response to an operator input), ease to provide a feed-back, tolerance to dynamic environments, ease to set up the system. Touch pads and head tracker qualifies quite well on all criteria, but devices such as DVI and eye trackers currenly present obvious limitations. Gesture and biopotentials have still serious limitations and uncertainties.

Characteristics / Technology	Fast	Reliable	Easy Feed-back	Dynamics Tolerance	Easy setup
Touch Screen	+	+	±	±	+
Touch Pad	+	+	+	+	+

Head-Tracker	+	+	+	+	+
Eye-Tracker	±	±	±	+	-
DVI	-	±	±	±	+
Gesture	±	-	-	±	-
EMG	±	±	-	+	-
EEG	± ?	±	-	±	-

Table 2: Compliance of control modes with various system and environmental criteria (+: good, ±: acceptable, -: non-acceptable, ?: questionable)

Table 3 examines the same technologies against some operator's usage criteria.

Comparison of touch-pads and touch screen show a clear advantage in favor of touch screens on these criteria. Actually there is a good complementarity between these two devices, also existing with accuracy (not considered here) suggesting that they could be used to create an heterogeneous redundancy. Head tracker appears to comply quite well to these criteria, while eye-tracker an DVI exhibit some uncertainties or weakness. Again, gesture and biopotentials are the least compliant modalities.

Characteristics / Technology	Low attention on control	High confidence	Intuitive	Easy error correction	Delay intent/input
Touch Screen	+	+	+	+	±
Touch Pad	±	+	±	±	±
Head-Tracker	+	+	+	+	±
Eye-Tracker	+	?	+	± ?	+
DVI	±	±	±	±	-
Gesture	±	±	±	- ?	± ?
EMG	-	-	±	- ?	+
EEG	-	-	±	- ?	+ ?

Table 3: Compliance of control modes with various oprator's usage criteria (+: good, ±: acceptable, -: non-acceptable, ?: questionable)

The same approach can also be applied to conventional manual controls, as grip-stop inceptors and small joysticks. Usually this kind of control qualifies individually very well against most of both engineering and operator's criteria, showing the robustness and efficacy of manual control solutions. It has to be remembered, however, that it is more the accumulation of these different controls in the control suite than their individual characteristics which creates difficulties relatively to short term memory management. Other cognitive criteria should be introduced to give an account of this kind of phenomenon and help to identify potential benefits of the new technologies, as alternative solutions or substitutes of manual control.

3.2 Cockpit simplification

True cockpit layout and functional simplification is only expected to apply to new cockpit design. Identifying potential areas of benefits offered by Alternative Control Technologies in this context becomes, therefore, very speculative.

An attractive goal could be to replace most dedicated control panels by non-conventional controls. Benefits could then be found not only on the functional side, with a considerably increase flexibility and ease of use, but also in regard of costs of installation, integration and maintenance. This is probably achievable at relatively short term with acceptably low risk, mainly using technologies such as touch screen, touch pad and DVI..

Another interesting area for application of Alternative Control Technology is linked with the introduction of very large reconfigurable displays in the cockpit. Such displays have been advocated for many years now and technological solutions, stimulated by general public applications, start to appear. On the display side, .some of. the benefits expected would be:

- Increased flexibility, as the displays windows size or location could rapidly be reconfigured following mission type, phase or even user experience or cognitive style. On request, most adapted size for a given situation could be rapidly obtained, from full screen to iconic
- Capacity to display complex tactical situations with an always appropriate size and resolution
- Allow Inflight mission planning and rehearsal.

Although, interacting with this kind of display to fully exploit its capability would necessitate very intuitive means to access the information displayed and reconfigure it following user's and context needs. Classical manual controls as joysticks or even touch-pads have already been shown to be quite poorly adapted to navigation in the various displays sub-divisions of such large displays. Use of interaction media, as eye-trackers, would be then of considerable interest for an intuitive localization and designation of the current point of interest. It is quite likely that interactions based on complementarity of modalities could be then envisioned with this kind of design. Some work is currently underway in this area.

A step further, probably in the long term, virtual cockpit could introduce huge benefits in terms of cockpit design, as it would not be anymore necessary to install physical head-down displays in the cockpit. Based on very large field of view

HMDs, this concept would use both implicit and explicit eye and head controls to interact with the system. It would represents an ideal field of application for other alternative control technologies, which could then contribute to the realization of a « joint cognitive system », closely associating the operator and the machine. This concept suppose, however, very significant progress to be achieved in various aspects of Alternative Control Technology and HMD, making it a high risk/ high pay-off option. Assuming acceptably low cost could be realized on this type of system, virtual crew station could also represent a highly portable and flexible solution for UAVs control stations

3.3 training considerations

A strong point to the introduction of Alternative Control Technologies is that they are supposed to be a lot more intuitive than conventional controls. This should imply that training needs would be reduced, yielding serious benefits as training is inevitably associated with costs. Things, however, may, not be as simple, as technology does not perfectly mediates the natural modalities used by the operator. On the other hand, in regard of memory management, using alternative control could help the pilot to reach more rapidly a given level of global proficiency on the aircraft system.

3.3.1 training on the control modality

It has been shown repeatedly with various Alternative Control technologies that experienced users would perform significantly better than naive ones, even on very easy tasks. It has also been reported that performance of speakers exposed to G-loads were improving from the beginning to the end of centrifuge experiments, probably as they learn to breathe and talk under G. The relationship between characteristics of technology and training issues are not very well understood. Some technologies, in their current status, explicitly call for some kind of training, as EMG and EEG. Gesture is also highly susceptible to require a substantial training need if communication capabilities are used.

3.3.2 impact on general training needs

This domain is very far to be clearly defined and, apparently, little work has been devoted to the impact of non-conventional controls on training. It could be expected that redundancy and alternative solutions could globally facilitate training on complex systems, as the operator's limited resources would be better used. This kind of issues definitely deserve some attention, since demonstration of training process improvements may constitutes a strong point for integration of Alternative Control Technologies in existing and future cockpit.

4. CONCLUSIONS

The current lecture has tried to review synthetically the various issues associated with the implementation of Alternative Control Technology in the aerospace environment. Most of the data presented in this lecture and preceding ones have been gathered through the activities of AGARD Working Group 25. Though oriented towards aerospace domain, such data may apply to other defense or even civilian applications. A comprehensive state of the art review has been conducted relatively to the different technological areas to be covered. Integration issues were approached following two converging pathways: human factors, including tools and design methodology considerations, human engineering and technical issues. Needs for future research and improvements were identified and efforts were devoted to assess benefits and challenges expected from the introduction of these new technologies in the cockpit.

To allow pilots the full benefits of Alternative Control technology, a noticeable amount of work remains to be done by researchers and engineers both in the human factors and engineering domains. Integration of these technology requires more than putting boxes side by side and physical connections to the aircraft system.. Similarly. to automation, problems to be solved have little chance to be resolved by a clumsy integration. Achieving a meaningful and smart implementation of these technologies will require a synergistic effort involving research labs, airframe and system manufacturers and equipment makers.

There is still a long way to go from the current « all manual status » to the « joint cognitive systems ». Hopefully, progress in the field of cognitive ergonomics and Alternative Control Technology will contribute to achieve this goal. Successive significant steps should be observed, alternative solutions, Large Interactive Displays and Virtual cockpit. Finally, the development of true « though-based » systems may allow one day to realize one of the oldest dream of humanity, already existing in Greek mythology, flying like a bird.

ACKNOWLEDGEMENT

The author wishes to thank all members of AGARD Working Group 25 for the dedication and commitment to the field of Alternative Control Technologies.

BIBLIOGRAPHY

Selected references to the literature on alternative control technology.

Prepared by Members of AGARD WG 25.

Canada	Dr. Bernard Hudgins, University of New Brunswick, Fredericton, New Brunswick
France	Dr. Alain Leger, SEXTANT Avionique, Saint-Médard-en-Jalles
	Mr. Dominique Pastor, SEXTANT Avionique, Saint-Médard-en-Jalles
	Dr. Pierre Dauchy, IMASSA - CERMA, Institut de Médicine Aérospatiale, Brétigny-sur-Orge
Germany	Dr. Hans Pongratz, Flugmedizinisches Institut der Luftwaffe, Fürstenfeldbruck
United Kingdom	Dr. Graham Rood, DERA, Farnborough
	Dr. Don Jarrett, DERA, Farnborough
	Mr. Allan South, DERA, Farnborough
	Dr. Karen Carr, British Aerospace, Sowerby Research Center, Bristol
United States	Dr. Grant McMillan, Air Force Research Laboratory, Wright-Patterson Air Force Base OH
	Dr. Timothy Anderson, Air Force Research Laboratory, Wright-Patterson Air Force Base OH
	Mr. Joshua Borah, Applied Science Laboratories, Bedford MA

1. OPERATIONAL RATIONALE

Advanced Aircraft Interfaces: The Machine Side of the Man Machine Interface AGARD Conference Proceedings 521 (AGARD-CP-521), 1992

Aviation Week 16 October 1995

Combat Automation for Airborne Weapon Systems: Man Machine Interface Trends and Technologies AGARD Conference Proceedings 520 (AGARD-CP-520), 1993

FAA Human Factors Team Report on: The Interfaces between Flight Crews and Modern Flight Deck Systems June 1996

Flight Vehicle Integration Panel Working Group 21 Glass Cockpit Operational Effectiveness AGARDAR-349, 1996

Gooderson C.Y.et al "The Hand Anthropometry of Male and Female Military Personnel" Army Personnel Research Establishment Memorandum 82M510, 1982

The Man Machine Interface in Tactical Aircraft Design and Combat Automation. AGARD Conference Proceedings No 425 (AGARD-CP-425), 1987

White G. and Becket, P. "Increased Aircraft Survivability using Direct Voice Input" AGARD CP 1983

2. SPEECH

"Proceedings of the Fourth International Conference on Spoken Language Systems", 3-6 Oct., 1996, Philidelphia, PA.

"Workshop on Audio-Visual Speech Processing", 26-27 Sept., 1997, Rhodes, Greece.

Acero, A., and Stern, R. M., "Environmental robustness in automatic speech recognition", in "Proc. Intl. Conf. Acoust., Speech, Signal Processing", 1990, pp 849-852.

Acero, A., and Stern, R. M., "Robust speech recognition by normalization of the acoustic space", in "Proc. Int. Conf. on Acoust., Speech, and Signal Processing", April 1991, pp 893-896.

Arslan, L. M, and Hansen, J. H. L., "Improved HMM training and scoring strategies with application to accent classification", in "Proc. Int. Conf. on Acoust., Speech, and Signal Processing", May 1996, pp 598-601.

Arslan, L. M, and Hansen, J. H. L., "Language Accent Classification in American English", Speech Communication, Vol. 18(4), pp 353-367, June/July 1996.

Benson, P., and Vensko, G., "A Spoken Language Understanding System for Phraseology Training of Air Traffic Controllers," Speech Technology, Media Dimensions, New York, Vol. 5, No. 1, Oct./Nov. 1989, pp 64-69.

Boll, S. F., "Speech enhancement in the 1980s: Noise suppression with pattern matching", in Furui, S., and Sondhi, M. M., (Eds), "Advances in Speech Signal Processing", M. Dekker, New York, 1992, pp 309-325.

Bourlard, H. and Wellekens, C. J., "Links between Markov models and multi-layer perceptrons", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 12, 1990, pp 1167-1178.

Bregler, C., Omohundro, S. M., Shi, J., and Konig, Y., "Towards a Robust Speechreading Dialogue System", in Stork., D. G., and Hennecke, M. E., (Ed) "Speech Reading by Humans and Machines: Models, Systems, and Applications", Berlin, Germany, Springer-Verlag, 1996, pp 408-423, (ISBN 0 540 61264 5).

Brooke, N. M., and Petajan, E. D., "Seeing speech: Investigations into the synthesis and recognition of visible speech movements using automatic image processing and computer graphics", in "Proc. Int Conf. Speech Input and Output: Techniques and Applications", Science Education and Technology Division of the IEE, Mar. 1986, pp 104-109.

Castiglione, D., and Goldman, J., "Speech and the Space Station", Speech Technology, Media Dimensions, New York, Vol. 2, No. 3, Aug./Sept. 1984, pp 19-27.

Davis, S. and Mermelstein, P., "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-28, No. 4, August 1980, pp 357-366.

Deller, J. R., Proakis, J. G., and Hansen, J. H. L., "Discrete-Time Processing of Speech Signals", Englewood Cliffs, USA, Macmillan Publishing Company, 1993.

Diakoloukas, V., Digalakis, V., Neumeyer, L., and Kaja, J., "Development of a dialect-specific speech recognizers using adaptation methods", in "Proc. Int. Conf. on Acoust., Speech, and Signal Processing", April 1997, pp 1455-1458.

Ephraim, Y., "A Bayesian estimation approach for speech enhancement using hidden Markov models", IEEE Trans. Acoust., Speech, Signal Process., April 1992, 40(4), pp 725-735.

Ephraim, Y., "Gain-adapted hidden Markov models for recognition of clean and noisy speech", IEEE Trans. Acoust., Speech, Signal Process., 40(4), April 1992, pp 725-735.

Ephraim, Y., Malah, D., and Juang, B.-H., "On the application of hidden Markov models for enhancing noisy speech", IEEE Trans. Acoust., Speech, Signal Processing, 37(12), December 1989, pp 1846-1856.

Erell, A., and Weintraub, M., "Estimation using log-spectral distance criterion for noise-robust speech recognition," in "Proc. Intl. Conf. Acoust., Speech, Signal Processing", 1990, pp 853-856.

Finn, K. E. and Montgomery, A. A., "Automatic optically-based recognition of speech", Patt. Recogn. Lett., 8(3), 1988, pp 159-164.

Flanagan, F. L., "Speech Analysis, Synthesis and Perception", Springer Verlag, New York, 1972.

Forney, G. D., "The Viterbi Algorithm", Proc. IEEE, Vol. 61, 1973, pp 268-278.

Ghitza, O., "Auditory nerve representation as a basis for speech processing", in Furui, S., and Sondhi, M. M., (Eds), "Advances in Speech Signal Processing", M. Dekker, New York, 1992, pp 453-485.

Goldschen, A., "Continuous Automatic Speech Recognition by Lipreading", PhD thesis, The George Washington University, Washington, DC, 1993.

Gulli CH.., Pastor D. Léger A., Sandor P.B., Clere J.M., Crateau P., « G-load effects and efficient acoustic Parameters for robust speaker recognition » in AGARD-CP-521, « Advanced Aircraft Interface : The Machine Side of the Man-Machine Interface », 22-1 to 14 1992

Hansen, J. H. L., and Arslan, L. M, "Foreign accent classification using source generator based prosodic features", in "Proc. Int. Conf. on Acoust., Speech, and Signal Processing", May 1995, pp 836-839.

Hennecke, M. E., Stork, D. G., and Prasad, K. V., "Visionary Speech: Looking Ahead to Practical Speech Reading Systems", in Stork., D. G., and Hennecke, M. E., (Ed) "Speech Reading by Humans and Machines: Models, Systems, and Applications", Berlin, Germany, Springer-Verlag, 1996, pp 408-423 (ISBN 0 540 61264 5).

Hermansky, H., "Perceptual linear predictive (PLP) analysis of speech", J. Acoust. Soc. Am., 87(4), April 1990, pp 1738-1752.

Hermansky, H., personal communications, 1995.

Hirsch, H., Meyer, P., and Ruehl, H. W., "Improved speech recognition using high-pass filtering of subband envelopes", in "Proc. of Eurospeech 1991", September 1991, pp 413-416.

Hoskins, J. W., "Voice I/O in the Space Shuttle", Speech Technology, Media Dimensions, New York, Vol. 2, No. 3, Aug./Sept. 1984, pp 13-18.

Howard, J. D., "Flight testing of the AFTI/F-16 voice interactive avionics system", in "Proc. Military Speech Technology", 1987, pp 76-82.

Humphries, J. J., Woodland, P. C., and Pearce, D., "Using accent-specific pronunciation modelling for robust speech recognition", in "Proc. Int. Conf. On Spoken Language Systems", Oct. 1996, pp 623-626.

Hunt, M. J., and Lefèbvre, C., "A Comparison of Several Acoustic Representations for Speech Recognition with Degraded and Undegraded Speech", in "Proc. Int. Conf. on Acoust., Speech, and Signal Processing", 1989, pp 262-265.

Junqua, J.-C., Wakita, H., and Hermansky, H., "Evaluation and optimization of perceptually based ASR front end", IEEE Trans. Speech and Audio Process., 1(1), January 1993, pp 39-48.

Kudo, I., Nakama, T., Watanabe, T., and Kameyama, R., "Data collection of Japanese dialects and its influence into speech recognition", in "Proc. Int. Conf. On Spoken Language Systems", Oct. 1996, pp 308-311.

Kumpf, K., and King., R. W., "Automatic Accent Classification of Foreign Accent Australian English Speech", in "Proc. Int. Conf. On Spoken Language Systems", Oct. 1996, pp 4-7.

Lee, C., Lin, C.-H., and Juang, B.-H., "A study on speaker adaptation of the parameters of continuous density hidden Markov models", IEEE Trans., 39(4), April 1994, pp 806-814.

Lee, K.-F., "Context-dependent phonetic hidden Markov models for speaker-independent continuous speech recognition", IEEE Trans. Acoust., Speech, Signal Process., 38(4), April 1990, pp 599-609.

Lee, K.-F., "Large-Vocabulary Speaker-Independent Continuous Speech Recognition: the SPHINX system", PhD thesis, Carnegie-Mellon, 1988.

Lee, K.-F., Hon, H.-W., and Reddy, R., "An overview of the SPHINX speech recognition system", IEEE Trans. Acoust., Speech, Signal Process., 38(1), January 1990, pp 35-45.

Lemoine, C., " Recherche de traits acoustiques de la parole bruitée par Analyse Multi-Résolution ", Thèse de l'Université de Bordeaux I, 1998.

Lowerre, B. T., "The Harpy Speech Recognition System", Doctoral Thesis, Carnegie-Mellon University, Pittsburgh, PA, 1976.

Mase, K., and Pentland, A., "Automatic lipreading by computer", in "Proc. IEICE Image Understanding Symposium", Apr. 1989, pp 65-70.

Mase, K., and Pentland, A., "Lip reading: Automatic visual recognition of spoken words", Opt. Soc. Am. Topical Meeting on Machine Vision, June 1989, pp 1565-1570.

Massaro, D.W., "Speech perception by ear and eye", in B. Dodd and R. Campbell, (Eds) "Hearing by Eye: the Psychology of Lip-reading", Lawrence Erlbaum Associates, London, 1987, pp 33-83.

McGuinness, "Effects of Feedback Modality in an Airborne Voice Communications Task", RAE Technical report TR 87072, December 1987.

Murveit, H., Butzberger, J., and Weintraub, M., "Reduced channel dependence for speech recognition", in "Proc. of the DARPA Speech and Natural Language Workshop", Harriman, NY, February 1992, pp 280-284.

Neumyer, L. and Wientraub, M., "Robust speech recognition in noise using adaptation and mapping techniques", in "Proc. Int. Conf. on Acoust., Speech, and Signal Processing", 1995, pp 141-144.

Pallett, D., and Fiscus, J., "1996 Preliminary Broadcast News Tests", in "Proc. of DARPA Speech Recognition Workshop", Feb. 1997, Virginia.

Pastor, D., "Diagnostic sur Signaux quasi-stationnaires par Décomposition en Ondelettes Orthonormales et Détection de Coefficients Significatifs", Thèse de l'Université de Rennes I, 1997.

Pastor, D., and Gulli, C., "DIVA 5 Dialogue Vocal pour Aéronef: Performance in Simulated Aircraft Cockpit Environments", Joint ESCA-NATO/RSG10 Tutorial and Workshop: Applications of Speech Technology, 1993, Lautrach.

Pastor, D., and Gulli, C., "Improving Recognition Rate in Adverse Conditions by Detection and Noise Suppression", in "Proc. ESCA Workshop on Speech Recognition in Adverse Conditions", 1992, Cannes-Mandelieu.

Petajan, E. D., "Automatic lipreading to enhance speech recognition", in "Proc. IEEE Global Telecom. Conf.", Nov. 1984, pp 265-272.

Petajan, E. D., "Automatic Lipreading to Enhance Speech Recognition", PhD thesis, University of Illinois, 1984.

Rabiner, L., and Juang, B.-W., "Fundamentals of Speech Recognition", Englewood Cliffs, NJ, Prentice-Hall, 1993, (ISBN 0 13 015157 2).

Rabiner, L., Levinson, S., and Sondhi, M., "On the application of vector quantization and hidden Markov models to speaker-independent isolated word recognition", Bell Sys. Tech. Journal, 62, 1983, pp 1075-1105.

Rajasekaran, P. K. and Doddington, G., "Robust speech recognition: initial results and progress", in "Proc. of the DARPA Speech Recognition Workshop", Palo Alto, CA, February 1986, pp 73-80.

Reisberg, D., "Easy to hear but hard to understand: a lip-reading advantage with intact auditory stimuli", in B. Dodd and R. Campbell, (Eds) "Hearing by Eye: the Psychology of Lip-reading", Lawrence Erlbaum Associates, London, 1987, pp 97-113.

Renals, S., Morgan, N., Bourland, H., Cohen, M., and Franco, H., "Connectionist Probability Estimators in HMM Speech Recognition", IEEE Trans. On Speech and Audio Processing, Vol. 2, No. 1, Part II, 1994.

Salazar, G., "Voice Recognition Makes its Debut on the NASA STS-41 Mission," Speech Technology, Feb/March 1991.

Salisbury, M., and Chilcote, J., "Investigating Voice I/O for the Airborne Warning and Control System (AWACS)," Speech Technology, Media Dimensions, New York, Vol. 5, No. 1, Oct/Nov 1989, pp 50-55.

Scherer, K. R., "Voice, Stress and Emotion", in M. H. Appley, R. Trumbull (Eds), Dynamics of Stress, New York, 1986, pp 157-179.

Sejnowski, T. J., Yuhas, B. P., Goldstein, M. H., and Jenkins, R. E., "Combining visual and acoustic speech signals with a neural network improves intelligibility", in D. S. Touretzky, (Ed) "Advances in Neural Information Processing", Morgan Kaufman, 1990.

Silsbee, P. L., "Computer Lipreading for Improved Accuracy in Automatic Speech Recognition", PhD thesis, University of Texas at Austin, 1993.

Silsbee, P.L., and Su, Q., "Audiovisual Sensory Integration Using Hidden Markov Models", in Stork., D. G., and Hennecke, M. E., (Ed) "Speech Reading by Humans and Machines: Models, Systems, and Applications", Berlin, Germany, Springer-Verlag, 1996, pp 33349, (ISBN 0 540 61264 5).

Simpson, C. A., Coler, C. R., and Huff, E. M., "Human Factors of Voice I/O for Aircraft Cockpit Controls and Displays", in "Proc. Of the Workshop on Standardization for Speech I/O Technology", National Bureau of Standards, Gaithersburg, Md., March 1982.

Stanton, B. J., "Robust recognition of loud and Lombard speech in the fighter cockpit environment", Doctoral Thesis, Purdue University, West Lafayette, IN, 1988.

Stork, D. G., Wolff, G., and Levine, E., "Neural network lipreading system for improved speech recognition", in Intl. Joint Conf. on Neural Networks, 1992, pp 285-295.

Stork., D. G., and Hennecke, M. E., (Ed) "Speech Reading by Humans and Machines: Models, Systems, and Applications", Berlin, Germany, Springer-Verlag, 1996, (ISBN 0 540 61264 5).

Sumby, W. H. and Pollock, I., "Visual contribution to speech intelligibility in noise", J. Acoust. Soc. Am., Vol. 26, 1954, pp 212-215.

Summerfield, Q., "Some preliminaries to a comprehensive account of audiovisual speech perception", in B. Dodd and R. Campbell, (Eds) "Hearing by Eye: the Psychology of Lip-reading", Lawrence Erlbaum Associates, London, 1987, pp 3-51.

Teixeira, C., Trancoso, I., and Serralheiro, A., "Accent Classification", ", in "Proc. Int. Conf. On Spoken Language Systems", Oct. 1996, pp 577-580.

Tishby, N., "A dynamical systems approach to speech processing", in "Proc. Intl. Conf. Acoust., Speech, Signal Processing", 1990, pp 365-368.

Varga, A. P., and Moore, R. K., "Hidden markov model decomposition of speech and noise", in "Proc. Intl. Conf. Acoust., Speech, Signal Processing", 1990, pp 845-848.

Waibel, A., "Prosodic knowledge sources for word hypothesization in a continuous speech recognition system", in Waibel, A., and Lee, K.-F., (Eds), "Readings in Speech Recognition", Morgan Kaufmann. San Mateo, CA, 1990, pp 534-537.

Waibel, A., "Prosody and Speech Recognition", PhD thesis, Carnegie-Mellon, 1986.

Waibel, A., Hanazawa, T., Hinton, G., Shikano, K., and Lang, K., "Phoneme recognition using time-delay neural networks", in "Proc. Int. Conf. on Acoust., Speech, and Signal Processing", 1988, pp 99-102.

Wesfreid, E., and Wickerhauser, M. V., "Adapted Local trigonometric Transforms and Speech Processing", IEEE Transactions on Signal Processing, Vol. 41, No. 12, 1993.

Wickerhauser, M. V., "Adapted Wavelet Analysis form Theory to Software", A. K. Peters, 1994, Massachusetts.

Williams C.S., "Designing Digital Filters", Englewood Cliffs, Prentice-Hall Inc. 1986, (ISBN 0-13-201856-X 01).

Woods, W. A., "Language processing for speech understanding," in Waibel A., and Lee, K.-F., (Eds), Readings in Speech Recognition, Morgan Kaufman, San Mateo, CA, 1990, pp 519-533.

Young, S. R., Hauptmann, A. C., Ward, W. H., Smith, E. T., and Werner, P., "High-level knowledge sources in usable speech recognition systems", in Waibel, A., and Lee, K.-F., (Eds), "Readings in Speech Recognition", Morgan Kaufmann, San Mateo, CA, 1990, pp 538-549.

Yuhas, B. P., Goldstein, M. H., and Sejnowski, T. J., "Integration of acoustic and visual speech signals using neural networks", IEEE Commun. Mag., Nov. 1989, pp 65-71.

Yuhas, B. P., Goldstein, M. H., Sejnowski, T. J., and Jenkins, R. E., "Neural network models of sensory integration for improved vowel recognition", Proc. IEEE, 78(10), Oct. 1990, pp 1658-1668.

3. HEAD POINTING

"An introduction to Honeywell helmet-mounted displays", Avionics Division, Honeywell, 1977.

"Glass cockpit operational effectiveness", AGARD AR-349, 1996.

"Operator , Organizational, Direct Support and General Support Maintenance Manual", US Army Technical Manual TM 9-1270-212-14&P, July, 1981.

« 1998 TLVs and BEIs", American Conference of Governmental Industrial Hygenists, Cincinnati, OH, 1998, (ISBN 88-2417-23-2), p 142.

Adelstein, B. D., Johnston, E. R., and Ellis, S. R., "Dynamic Response of Electromagnetic spatial displacement trackers", Presence, 5, 3, 1996, pp 302-318.

Applewhite, H. L., "A new ultrasonic positioning principle yielding pseudo-absolute location", in Singh, G., Feiner, S., K., and Thalmann, D. (Eds.) "Virtual Reality Software and Technology. Proceedings of the VRST '94 Conference", Singapore, 1994, pp. 175-83.

Barnes, G. R. and Sommerville, G. P., "Visual target aquisition and tracking performance using a helmet-mounted sight", Aviation, Space, and Environmental Medicine, April, 1978, pp 565-572.

Bizzi, E., "Eye-head coordination". In Brooks, V. B. (Ed) "Handbook of Physiology, The Nervous System", Sect 1, Vol 2, Part 2, Ch29, Bethesda, MD, American Physiological Society,. 1981, pp 1321-1336.

Bizzi, E., Kalil, R. E., and Morasso, P., "Two modes of active eye-head coordination in monkeys", Brain Research, 40, 1972, pp 45-48.

Blood, E., "Device for Quantitatively Measuring the relative position and orientation of two bodies in the presence of metals utilizing DC magnetic fields", US patent no. 4,849692, 1989.

Blood, E., "Device for Quantitatively Measuring the relative position and orientation of two bodies in the presence of metals utilizing DC magnetic fields", US patent no. 4,945,305, 1990.

Borah, J., "Investigation of Eye and Head Controlled Cursor Positioning Techniques", US Air Force report AL/CF-SR-1995-0018, September 1995.

Brindle, J. H., "Advanced helmet tracking technology developments for naval aviation", in "SAFE association, 33d Annual Symposium", Reno, NV, Oct. 23-25, Proceedings (A96-1671603-54), 1995, pp 34-53.

Brown T. C., "AH-64 Apache night vision system" in "Night vision '92" Conference, London, 1992.

Church T. O. and Bennett W. S., "System automation and pilot-vehicle-interface for unconstrained low-altitude night attack" in "Combat automation for airborne weapon systems: Man-machine interface trends and technologies", Edinburgh, UK. AGARD-CP-520, 1992.

Durlach, N. J. and Mavor, A. S., "Virtual Reality Scientific and Technological Challenges", Washington, D.C., National Academy Press, 1995, pp188-204.

Emura, A. and Tachi, S., "Compensation of time lag between actual and virtual spaces by multi-sensor integration", in "Proceedings of the 1994 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems", Las Vegas, NV, 1994.

Fitts, P. M., "The information capacity of the human motor system in controlling the amplitude of movement", Journal of Experimental Psychology, 47, 1954, pp 381-391.

Fong, K. L., "Maximizing +Gz Tolerance in Pilots of High Performance Combat Aircraft", US Air Force Report AL-SR-1993-0001, December 1992.

Foxlin, E.; and Durlach, N., "An inertial head-orientation tracker<with automatic drift compensation for use with HMD's", in Singh, G., Feiner, S., K., and Thalmann, D. (Eds.) "Virtual Reality Software and Technology. Proceedings of the VRST '94 Conferrence", Singapore, 1994, pp 159-73.

Furness T. A. and Kocian D. F., "Putting humans in virtual space" The Society for Computer Simulation, Simulation Series, 16, 2, San Diego, CA, 1986, pp 214-230.

Furness T. A., "The effect of whole body vibration on the perception of the helmet-mounted display", Ph. D. dissertation, Univ. Southampton (unpublished), 1981.

Glanville, A. D., and Kreezer,G., "The maximum amplitude and velocity of joint movements in normal male human adults", Human Biology, 9, 1937, p 197.

Griffin, M. J., "Vertical vibration of seated subjects: Effects of posture, vibration level and frequency", Aviation, Space and Environmental Medicine, 46, 1975, pp 269-276.

Hericks, J., Parise, M., and Wier, J., "Breaking down the barriers of cockpit metal in magnetic head tracking", in "Proceedings of the SPIE Head Mounted Displays", Orlando, FL, April 8-10, 1996.

Hertzberg, H. T. E., "Human Anthropology" in VanCott, H. P. and Kinkade, R. G., (Eds) "Human Engineering Guide to Equipment Design", American Institutes for Research, Wahington D.C., 1972.

Jacobs R. S., Triggs T. J., and Aldrich J. W., "Helmet-mounted display/sight system study" US Air Force Technical Report AFFDL-TR-70-83 Vol 1, 1970.

Jagacinski, R .J., and Monk, D. L., "Fitts' law in two dimensions with hand and head movements", Journal of Motor behavior, 17, 1985, pp 77-95.

Jarrett, D. N., "Helmet position sensor and loading mechanism", DRA working paper DRA-FS-93-WP892, 1993.

Kaye M. G., Ineson J., Jarrett D. N. and Wickham G., "Evaluation of virtual cockpit concepts during simulated missions", in "Helmet-mounted displays II" SPIE 1290, 1990, pp 236-245.

Kocian, D. F., and Task, H. L., "Visually Coupled Systems Hardware and the Human Interface" In Barfield, W., and. Furness, T. A, (Eds) "Virtual Environments and Advanced Interface Design", New York, Oxford University Press, 1995.

Kuipers, J. B., "SPASYN -- an electromagnetic relative position and orientation tracking system", IEEE Transactions on Instrumentation and Measurement, IM-29, 4, 1980, pp 462-466.

Lee, J. M.; Chartier, V. L.; Hartmann, D. P.; Lee, G. E.; Pierce, K. S., "Electrical and Biological Effects of Transmission Lines: A Review", U.S. Department of Energy Report DOE/BPA-945, Jun 1989.

Leger A., Sandor P. "Désignation de cible sous facteur de charge: intérêt et limites du viseur de casque", AGARD-CP 478, AMP symposium on "Situational awareness in aerospace operations", Copenhagen, Denmark, .11, 2-5 October, 1989, pp 1-10.

Leger A., Sandor P., Bourse C., Alain A., "Réponse biomécanique de la tête aux accélérations +Gz: Intérêt pour les études en simulation de combat", AGARD CP-517, "Helmet Mounted Displays and Night Vision Goggles", Pensacola, Fl, 6,-1991, pp 1-9.

Leger A., Sandor P., Clere J. M., Ossard G., "Mobilité de la tête et facteur de charge: approche expérimentale en centrifugeuse", AGARD-CP 471, AMP Symposium on "Neck injury in advanced military aircraft environments", Munich, Germany, 1989.

Leger A., Sandor P., Troseille, X., "Désignation d'objectifs sous facteur de charge: poursuite de cibles mobile", R.E. N° 32 CEV/SE/LAMAS, 1990.

Leger A., unpublished observations, 1993.

Leigh, R. J., and Zee, D. S., "The Neurology of Eye Movements", Philadelphia, F. A Davis Company, 1983, pp 109-123.

Lewis C. H., Griffin M. J., "Predicting the effect of vibration frequency and axis and seating conditions on the reading of numeric displays", Ergonomics, 23, 1980, pp 485-507.

Lin, M. L., Radwin, R .G., and Vanderheiden, G .C., "Gain effects on performance using a head-controlled computer input device", Ergonomics, 35, 2, 1992, pp 159-175.

Lydick L. N., "Head-steered sensor flight test results and implications" "Combat automation for airborne weapon systems: Man-machine interface trends and technologies", Edinburgh, UK. AGARD-CP-520, 1992.

MacKenzie, I. S., "Fitts' law as a research and design tool in human-computer interaction", Human Computer Interaction, 7, 1992, pp 91-139.

Mourant. R. R. and Grimson, C. G., "Predictive head-movements during automobile mirror sampling", Perceptual and Motor Skills, 44,. 1977, pp 283-286.

Raab, F. H., Blood, E. B., Steiner, T. O., and Jones, H. R., "Magnetic Position and Orientation Tracking System", IEEE transactions on Aerospace and Electronic Systems, AES-15, 5, 1979.

Radwin, R. G., "A method for evaluating head-controlled computer input devices using Fitts' law", Human Factors, 32, 4, 1990, pp 423-438.

Rolwes M. S., "Design and flight testing of an electronic visibility system" in "Helmet-mounted displays II", SPIE 1290, 1990, pp 108-119.

Rowlands, G. F., "The transmission of vertical vibration to the head and shoulders of seated men", Royal Aircraft Establishment Technical Report TR-77068. Farnborough, England, 1977.

Sandor P. B., Leger A., "Tracking with a restricted field of view: performance and eye-head coordination aspects", Aviat. Space Environ. Med; 62, 11, 1991, pp 1026-31.

Spitz, G., "Target acquisition performance using a head mounted cursor control device and a stylus with digitizing tablet", in "Proceedings of the Human Factors Society 34th Annual Meeting", 1990, pp 405-409.

Sutherland, I. E., "A head mounted 3 dimensional display", in "1968 Fall Joint Computer Conference, AFIPS Conference Proceedings", 33, 1968, pp 757-764.

Tatham, N. O., "The effects of turbulence on helmet-mounted sight accuracies", AGARD CPP 267, 1979.

Viviani P., Berthoz A., "Dynamics of the head-neck system in response to small perturbations: Analysis and modeling in the frequency domain", Biol. Cybernetics 19, 1975, pp 19-37.

Walsh, M. L., and Donnally, K. E., "Power frequency electric and magnetic field exposure and human health", in "35th Cement Indus. Tech. Conf", Toronto, Canada, May 23-27, IEEE Cat. No. 93CH3268-0, 1993, pp 279-88.

Welford, A. T., "The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress", Ergonomics, 3, 1960, pp 189-230.

Wells, M. J. and Griffin, M. J., "A review and investigation of aiming and tracking performance with head-mounted sights", IEEE Trans on Systems, Man and Cybernetics, SMC-17, 2, 1987, pp 210-221.

4. EYE POINTING

Allison, R. S., Eizenman, M. and Cheung, B. S. K., "Combined head and eye tracking system for dynamic testing of the vestibular system", IEEE Trans on Biomedical Engineering, Vol. 43, No 11, November 1996

Bahill, A. T., Adler, D. and Stark, L. "Most naturally occurring human saccades have magnitudes of 15 degrees or less", Investigative Ophthalmology, 14, 1975, pp 468-469.

Barnes, G. R., Benson, A. J. and Prior, A. R. J., "Visual-vestibular interaction in the control of eye movement", Aviation, Space and Environmental Medicine, 49, 1978, pp557-564.

Bergin, J. R. and Julez, B., "Parallel versus serial processing in rapid pattern discrimination" Nature, 303, pp 696-698.

Bolt, R. A., "The Human Interface: Where People and Computers Meet", Lifetime Learning Publications, London, UK. 1984.

Borah, J., "Helmet Mounted Eye Tracking for Virtual Panoramic Display Systems - Volume II: Eye Tracker Specification and Design Approach", US Air Force report AAMRL-TR-89019, August 1989.

Borah, J., "Helmet Mounted Eye Tracking for Virtual Panoramic Display Systems - Volume I: Review of Current Eye Movement Measurement Technology", US Air Force report AAMRL-TR-89019, 1989

Borah, J., "Investigation of Eye and Head Controlled Cursor Positioning Techniques", US Air Force report AL/CF-SR-1995-0018, September 1995.

Brennan, D. H., "Vision and visual protection in fast jet aircraft", in "Visual effects in the high performance aircraft cockpit", AGARD LS-156, 1988

Brinicombe, A. M., Boyce, J. F. and Durnell, L., "Direction of regard determination", in Delogne, P., (Ed) "Proc. Intl. Conf. on Image Processing", Lausanne, Switzerland. IEEE Signal Processing Society, 1996.

Calhoun, G. L. and Janson, W. P., "Eye control interface considerations for aircrew station design", in "Sixth European Conference on Eye Movements", Leuven, Belgium, 1991.

Calhoun, G. L. and Janson, W. P., "Eye Line-of-Sight Control Compared to Manual Selection of Discrete Switches", US Air Force report AL-TR-1991-0015, NTIS: AD-A273 019, 1991.

Calhoun, G. L., Janson, W. P. and Arbak, C. J., "Use of eye control to select switches", in "Proceedings of the Human Factors Society - 30th Annual Meeting", 1986, pp 154-158.

Card, S., Moran, T. and Newell, A., "The psychology of Human Computer Interaction", Hillsdale, Lawrence Erlbaum Associates, 1983.

Co-ordinate systems for describing eye movements. Section 1.903 in Boff, K. R. and Lincoln J. E., (Eds) "Engineering Data Compendium, Human Perception and Performance", US Air Force A.A.M.R.L., Ohio 1988.

Delabarre E. B., "A method for recording eye-movements", Amer. J. Psychol., 9, 572, 1898.

Ferman L., Collewijn H., Jansen T. C. and Van den Berg A. V., "Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral coil technique", Vision research, 27, 1987. pp 818-828

Fong, K. L., "Maximizing +Gz Tolerance in Pilots of High Performance Combat Aircraft, Interim Report", US Air Force report AL-SR-1993-0001, December 1992.

Frecher, R. C., Eizenman, M., and Hallet, P. E., "High precision real-time measurement of eye position using the first Purkinje image" in Gale, A.G. and Johmson, F., (Eds) "Theoretical and applied Aspects of Eye Movement Research", North-Holland, Elsevier Science Publishers B. V., 1984.

Glenn, F. A. Harrington, N., Iavecchia, H. P., and Stokes, J., "An Oculometer and Automated Speech Interface System", in "Analytics, Technical Report 1920", Analytics, Willow Grove, PA, May 1984.

Griffin, M. J., "Vertical vibration of seated subjects: Effects of posture, vibration level and frequency", Aviation, Space and Environmental medicine, 46, 1975, pp269-276.

Hallet,P.E., "Eye Movements", in Boff, K. R., Kaufman, L., and Thomas, M. P. (Eds) "Handbook of Perception and Human Performance - Vol. 1", New York, John Wiley and Sons. 1986.

Hatfield, F., Jenkins, E. and Jennings, M. W., "Eye/Voice Mission Planning Interface (EVMPI)", US Air Force report TR-J103-1, 1995.

Jacob, R. K., "Eye Tracking in Advanced Interface Design", in Barfield, W. and Furness, T. A. (Eds.) "Virtual Environments and Advanced Interface Design", New York, Oxford University Press, 1995.

Jarrett, D. N., "Helmet-mounted devices in low flying high speed aircraft", AGARD CPP 267, 1979.

John, B. E. and Kieras, D. E., "The GOMS Family of Analysis Techniques: Tools for Design and Evaluation", CMU Technical Report CMU-CS-94-181, Carnegie-Mellon University, August 1994.

Julez, B., Gilbert, E. N., Shepp, L. A., and Risch, H. L., "Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited", Perception, 2, 1973, pp 391-404.

Kenyon, R. V., Zeevi, Y.Y., Wetzel, P. A., and Young, L. R., "Eye movement in response to single and multiple targets", US Air Force report AFHRL-TR-84-29, 1985.

Kliegle, R. and Olson, R. K., "Reduction and calibration of eye monitor data", Behavior Research Methods & Instrumentation, 13, 2, 1981, pp 107-111.

Kocian, D. F. and Task, H. L., "Visually Coupled Systems Hardware and the Human Interface", in Barfield, W., and Furness, T. A., (Eds) "Virtual Environments and Advanced Interface Design", New York, Oxford University Press, 1995.

Koons, D. B., Sparrell, C. J. and Thorisson, K. R., "Integrating Simultaneous Input from Speech, Gaze, and Hand Gestures," in Maybury, M. T. (Ed.) "Intelligent Multimedia Interfaces", Menlo Park, AAI Press/The MIT Press, 1993.

Kowler, E. and McKee, S., "Sensitivity of smooth eye movement to small differences in target velocity", Vision Research, 27, 6, 1987, pp 993-1015.

Lambert, R. H., Monty, R. A. and Hall, R. J., "High-speed processing and unobtrusive monitoring of eye movements", Behavior Research Methods and Instrumentation, 6, 1974, pp 525-530.

Marx E. and Trendelenburg W., "Uber die genauigkeit der einstellung des auge biem fixieren", Z. Sinnesphysiol. 45, 1911, pp 87-102.

McConkie, G. W., "Evaluating and reporting data quality in eye movement research", Behaviu Research Methods & Instrumentation, 13, 2, 1981, pp 97-106.

Michael, J. A., Melvill Jones, G., "Dependence of visual tracking capability upon stimulus predictability", Vision Research, 6, 1966, pp 707-716.

Nodine, C. F., Kundel, H. L., Toto, L. C., Krupinski, E. A., "Recording and analysing eye-position data using a microcomputer workstation", Behavior Research Methods, Instruments & Computers, 24, 1992, pp 475-485.

Orschansky J., "Eine methode die augenbewegungen direkt zuuntersuchen (opthalmographie)", Zbl. Physiol., 12, 785, 1898.

Peli, E. and Zeevi, Y. Y., "Multiple visual feedback loops in eye movement control". in "XII International Conference on Medical and Biological Engineering", Jerusalem, 1979.

Robinson D. A., "A method for measuring eye movement using a scleral search coil in a magnetic field", IEEE Transactions on Biomedical Electronics, BME-10, 1963, pp 137-145.

Rowlands, G. F., "The transmission of vertical vibration to the head and shoulders of seated men", Royal Aircraft Establishment Technical Report TR-77068, 1977.

Sandor, P. B., Hortolland, I., Poux, F., and Leger, A., "Orientation du regard sous facteur de Charge Aspects methodologiques Resultats preliminaires", in "AGARD Meeting on Virtual Interfaces: Research and Applications", October, 1993.

Scinto, L. F. M., "Retinal inhomogeneity and the allocation of focal attention during fixation", in "The Annual Meeting of the Applied Vision Association" St. John's College, Oxford. 1988.

Shackel B., "Eye movement recording by electro-oculography", in "A manual of Psychophysiological methods", North-Holland Publ. Co., 1967.

Sheena, D. and Borah, J., "Compensation for some second order effects to improve eye position measurements", in Fisher, D. F., Monty, R. A., and Senders J. W., (Eds) "Eye Movements: Cognition and Visual Perception", Hillsdale, Lawrence Erlbaum Associates, 1981.

Sliney, D. H. and Wolbarsht, M., "Safety with Lasers and Other Optical Sources: A Comprehensive Handbook", Hew York, Plenum Pres, 1980.

St.Cyr G. L. and Fender, D. H., "Non-linearities of the human oculomotor system: Gain", Vision Research, 9, 1969, pp 1235-1246.

Starker, I., and Bolt, R. A., "A gaze-response self-disclosing display", in "Proceeding of the ACM CHI '90 Human Factors in Computing Systems Conference", New York, Addison Wesley/ACM Press, 1990, pp 3-9.

Steinman, R. M. and Collewijn, H., "Binocular Retinal Image Motion During Active Head Rotation", Vision Research, 20, 1980, pp 415-429.

Takeda, T., Fukui, Y., Ikeda, K. and Iide, T., "Three-dimensional optometer III", Applied Optics, 32, 22, 1993, pp 4155-68.

Tatham, N. O., "The effects of turbulence on helmet-mounted sight accuracies", AGARD CPP 267, 1979.

Viveash J. P., Belyavin A. J., "Eye movements under operational conditions", in Waters M. and Stott J. R. R. (Eds) "Journal of Defence Science", 1, 2, 1996.

Ware, C. and Mikaelian, H. T., "An evaluation of an eye tracker as a device for computer input", in by Carroll, J.M. and Tanner, P. P., (Eds) "Proceedings of Human Factors in Computing Systems and Graphics Interface Conference", Toronto, Canada. 1987, pp. 183-188.

Wells, M. J. and Griffin, M. J., "A review and investigation of aiming and tracking performance with head-mounted sights", IEEE Trans on Systems, Man and Cybernetics, T-SMC/17, 2, 1987, p12094.

Yamada. M., "Head and eye coordination analysis and a new gaze analyzer developed for this purpose", in d'Ydewalle G. and Van Rensbergen, J. (Eds) "Visual and oculomotor functions", Elsevier, 1994, pp 423-434.

Yarbus A. L., "Eye Movements and Vision", New York, Plenum Press, 1967.

Young L. R. and Sheena D., "Eye Movement Measurement Techniques", in Webster (Ed) "Encyclopedia of Medical Devices and Instrumentation", New York, John Wiley & Sons, 1988

Young, L.R., "The sampled data model and foveal dead zone for saccades", in Zuber B. L. (Ed) "Models of Oculomotor Behavior and Control", Boca Raton, CRC Press, 1981.

5. GESTURE

"Progress in Gestural Interaction", Proceedings of Gesture WorkshopÊ'96, University of York, March 1996.

Baudel, T., and Beaudoin-Lafon, M., "Charade: Remote control of objects using free-hand gestures", Communications of the Association for Computing Machinery, pp 28-35, July 1993.

Brooks, F. P., "Grasping reality through illusion: Interactive graphics serving science". In "Human Factors in Computing Systems", Proceedings of CHI '88 -, ACM, New York, 1988, pp 1-11.

Brooks, F. P., Ouh-Young, M., Batter, J. J., and Kilpatrick, J., "Project GROPE: Haptic displays for scientific visualisation", Computer Graphics, 24, 4, August 1990.

Burdea, G., and Coiffet, P., "La réalité virtuelle", France, Hermès, 1993 (ISBN 2-86601-386-7).

Buxton, B., "A directory of sources for input technologies", 1998, available on the Web as http://www.dgp.utoronto.ca/people/BillBuxton/InputSources.html

Cadoz, C., "Le geste canal de communication homme/machine, la communication 'instrumentale'", Technique et science informatiques, 13(1), 1994, pp 31-61.

Davis, J., and Shah, M., "Gesture recognition", University of Central Florida Technical Report CS-TR-93-11, 1993.

Dix, A., Finlay, J., Abowd, G., and Beale, R., "Human-Computer Interaction", UK, Prentice-Hall, 1993 (ISBN 0-13-437211-5).

Essa, I. A. and Pentland, A. P., "Coding, analysis, and recognition of facial expressions", Technical Report No. 325, MIT Media Lab, April 1995.

Fukumoto, M., Mase, K., and Suenaga, Y., "Real-time detection of pointing actions for a glove-free interface", in IAPR Workshop on Machine Vision Applications, December 7-9, 1992, pp 473-476.

Goble, J. R., Suarez, P. F., Rogers, S., K., Ruck, D. W., Arndt, C., and Kabrisky, M., "A facial feature communications interface for the non-verbal", IEEE Trans. On Engineering in Medicine and Biology, Sept 1993.

Hofmann, F. G., and Hommer, G., "Analyzing Human Gestural Motions using Acceleration Sensors", in "Progress in Gestural Interaction", Proceedings of Gesture WorkshopÊ'96, University of York, March 1996.

Hunke, M. and Waibel, A., "Face locating and tracking for human computer interaction", in 28th Asilomar Conference on Signals, Systems, and Computers, Monterey, CA, Nov. 1994.

Ineson, J., and Parker, C. C., "The accuracy of virtual touch", DRA Working Paper DRA/AS/MMI/WP95036/1, 1995.

Ineson, J., Parker, C. C., and Evans, A., "A comparison of head-out and head-in selection mechanisms during simulated flight", DERA Customer Report DERA/AS/SID/510/CR97153, 1997.

Kamp, J.-F., and Poirier, F., "Un dispositif tactile pour la commande en véhicule : étude d'une utilisation sans retour visuel", in "9èmes journées sur l'ingénierie de l'interaction Homme-Machine (IHM'97)", 1997 (ISBN 2-85428-459-3).

Krueger, M., "Artificial Reality", 2nd edition, Addison-Wesley, 1991.

Kurtenbach, G., and Buxton, W., "The Limits of Expert Performance Using Hierarchic Marking Menus", in ACM and IFIP joint conference on Human Factors in Computing Systems (INTERCHI'93), pp 482-487, 1993.

Limantour, P., "Medialab: Masters of Motion Capture", Computer Graphics World, October 1996.

Negroponte, N., "Being digital", Coronet, Hodder & Stoughton, 1995 (ISBN 0-340-64930-5)

Pritzwill, S., Leven, R., Zienert, H., Hanke, T., Henning, J., et al., "HamNoSys (version 2.0) Ñ Hamburg Notation System for Sign Languages / An introductory guide", in International Studies on Sign Language and the Communication of the Deaf, vol. 5, Hamburg, Germany, 1989.

Reising, J. M., et al., "New cockpit technology: unique opportunities for the pilot", SPIE conference, Orlando, April 1994.

Reising, J. M., Liggett, K. K., and Hartsock, D. C., "Exploring techniques for target designation using 3-D stereo map displays", International Journal of Aviation Psychology, 3(3), pp 169-187.

Rowley, H. A., Baluya, S., and Kanade, T., "Human face detection in visual scenes", Technical Report CMU-CS-95-158, CS Department, CMU, 1995.

Rubine, D., "The automatic recognition of gestures", Ph. D. thesis, Carnegie-Mellon University, 1991.

Shneiderman, B., "Designing the User Interface — Strategies for Effective Human-Computer Interaction", 2nd edition, Addison-Wesley, 1992 (ISBN 0-201-57286-9).

So, R. H. Y., "Comparison of the transmission of vertical seat vibration to the head and finger in a stationary target aiming task", United Kingdom and French Joint Meeting on Human Response to Vibration, 1988.

Solz, T. J., et al., "3-D stereo displays: how to move in the third dimension", SPIE conference, Orlando, May 1995.

Solz, T. J., et al., "The use of aiding techniques and varying depth volumes to designate targets in 3-D space", Proc. 38th Human Factors Soc. AGM, 1994.

Sturman, D. J., and Zeltzer, D., "A survey of glove-based input", IEEE Computer graphics and applications, 14, 1, January 1994, pp 30-39.

Sung, K. and Poggio, T., "Example-based learning for view-based human face detection", Technical Report 1521, MIT AL Lab., 1994.

Takemura, H., Tomono, A., and Kobayashi, Y., "A study of human-computer interaction via stereoscopic display", in "Work with Computers: Organisational, Management, Stress and Health Aspects", pp 496-503, M. J. Smith and G. Salvendy (eds), Elsevier Science Publishers, 1989.

Ward, M., Azuma, R., Bennett, R., Gottschalk, S., and Fuchs, H., "A demonstrated optical tracker with scalable work area for head-mounted display systems", in "1992 Symposium on interactive 3D graphics", Association for Computing Machinery, Cambridge, 1992, pp 43-52.

Wellner, P., "Interacting with Paper on the DigitalDesk", Communications of the Association for Computing Machinery, 36, 7, July 1993, pp 87-96.

White, J. L., et al. "Virtual cockpit concepts: an evaluation of data entry techniques", DRA Report DRA/AS/MMI/CR95168, 1995.

Yacoob, Y. and Davis, L. S., "Recognizing Facial Expression", Technical Report CS-TR-3265, University of Maryland, Computer Vision Laboratory, May 1994.

6. BIOPOTENTIALS

Barreto, A.B., Taberner, A.M., and Vicente, L.M., "Classification of Spatio-Temporal EEG Readiness Potentials towards the Development of a Brain-Computer Interface", in "Proceedings of the 1996 IEEE SouthEastcon Conference", IEEE, 1996, pp 100-103.

Calhoun, G.L., and McMillan, G.R., "EEG-Based Control for Human-Computer Interaction", in "Proceedings of the Third Annual Symposium on Human Interaction with Complex Systems", IEEE Computer Society Press, 1996, pp 4-9.

Calhoun, G.L., McMillan, G.R., Morton, P.E., Middendorf, M.S., Schnurer, J.H., Ingle, D.F., Glaser, R.M., and Figoni, S.F., "Functional Electrical Stimulator Control with a Direct Brain Interface", in "Proceedings of the RESNA 18th Annual Conference", 1995, pp 696-698.

Chatrian, G.E., Petersen, M.C., and Lazarete, J.A., "The Blocking of the Rolandic Wicket Rhythm and Some Central Changes Related to Movement", Electroencephalography and Clinical Neurophysiology, 11, 1959, pp 497-510.

Childress, D., "A Myoelectric Three State Controller Using Rate Sensitivity", in "Proceedings of ACEMB", 1969, pp S4-S5.

Clark, J.E., and Phillips, S.J., "The Efficacy of Using Human Myoelectric Signals to Control the Limbs of Robots in Space", NASA-CR-182901, 1988.

DeLuca, C.J., "Myoelectric Manifestations of Localized Muscular Fatigue in Humans", CRC Critical Reviews in Biomedical Engineering, 11, 1984, pp 251-279.

Donchin, E., Karis, D., Bashore, T.R., Coles, M.G.M., and Gratton, G., "Cognitive Psychophysiology and Human Information Processing., in Coles, M.G.H., Donchin, E., and Porges, S.W., (Eds) "Psychophysiology: Systems, Processes, and Applications", New York, Guilford Press, 1986.

Dorcas, D., and Scott, R.N., "A Three State Myoelectric Control", Medical Biological Engineering, 4, 1966, pp 367-372.

Dr. Lisa Dolev, Personal Communication, 1996.

Farry, K.A., Walker, I.D., and Baraniuk, R.G., "Myoelectric Teleoperation of a Complex Robotic Hand", IEEE Transactions on Robotics and Automation, 12, 5, 1996, pp 775-778.

Farwell, L.A., and Donchin, E., "Talking Off the Top of Your Head: Toward a Mental Prosthesis Utilizing Event-Related Brain Potentials", Electroencephalography and Clinical Neurophysiology, 70, 1988, pp 510-523.

Fernandez, J., Farry, K.A., and Cheatham, J.B., "Waveform Recognition Using Genetic Programming: The Myoelectric Signal Recognition Problem", Genetic Programming 1996 Conference, 1996.

Gevins, A., Leong, H., Du, R., Smith, M.E., Le, J., DuRousseau, D., Zhang, J., and Libove, J., "Towards Measurement of Brain Function in Operational Environments", Biological Psychology, 40, 1995, pp 169-186.

Gevins, A.S., Bressler, S.L., Cutillo, B.A., Illes, J., Fowler-White, R.M., Miller, J., Stern, J., Jex, H., "Effects of Prolonged Mental Work on Functional Brain Topography", Electroencephalography and Clinical Neurophysiology, 76, 1990, pp 339-350.

Gevins, A.S., Morgan, N.H., Bressler, S.L., Cutillo, B.A., White, R.M., Illes, J., Greer, D.S., Doyle, J.C., and Zeitlin, G.M., "Human Neuroelectric Patterns Predict Performance Accuracy", Science, 235, 1987, pp 580-585.

Graupe, D., Salahi, J., and Kohn, K.H., "Multifunction Prosthesis and Orthosis Control via Microcomputer Identification of Temporal Pattern Differences in Single-Site Myoelectric Signals", IEEE Transactions on Biomedical Engineering, BME-29, 4, 1982, pp 17-22.

Harms-Ringdahl, K., Ekholm, J., Schuldt, K., Linder, J., and Ericson, M., "Assessment of Jet Pilots' Upper Trapezius Load Calibrated to Maximal Voluntary Contraction and a Standardized Load", Journal of Electromyography and Kinesiology, 6, 1, 1996, pp 67-72.

Hiraiwa, A., Shimohara, K., and Tokunaga, Y., "EEG Topography Recognition by Neural Networks", IEEE Engineering in Medicine and Biology, 19, 1990, pp 39-42.

Hogan, N., and Mann, R.W., "Myoeletric Signal Processing: Optimal Estimation Applied to Electromyography - Part 1: Derivation of the Optimal Myoprocessor", IEEE Transactions on Biomedical Engineering, BME-27, 7, 1980, pp 282-295.

Hudgins, B., Parker, P., and Scott, R.N., "A New Approach to Multifunction Myoelectric Control", IEEE Transactions on Biomedical Engineering, BME-40, 4, 1993, pp 82-94.

Hudgins, B., Parker, P., and Scott, R.N., "Control of Artificial Limbs Using Myoelectric Pattern Recognition", Medical and Life Sciences Engineering, 13, 1994, pp 21-38.

Humphrey, D., and Kramer, A.F., "Toward a Psychophysiological Assessment of Dynamic Changes in Mental Workload", Human Factors, 36, 1994, pp 3-26.

Jacobsen, S., Knutti, D., Johnson, R., and Sears, H., "Development of the Utah Arm", IEEE Transactions on Biomedical Engineering, BME-29, 4, 1982, pp 249-269.

Jasper, H.H., and Penfield, W., "Electrocorticograms in Man: Effect of the Voluntary Movement upon the Electrical Activity of the Precentral Gyrus", Arch. Psychiat. Z. Neurol., 183, 1949, pp 163-174.

Jung, T-P., Makeig, S., Stensmo, M., and Sejnowski, T.J., "Estimating Alertness from the EEG Power Spectrum", IEEE Transactions on Biomedical Engineering, 44, 1, 1997, pp 60-69.

Junker, A., Berg, C., Schneider, P., and McMillan, G.R., "Evaluation of the CyberLink Interface as an Alternative Human Operator Controller", US Air Force Technical Report AL/CF-TR-1995-0011, 1995.

Kang, W., Cheng, C., Lai, J., Shiu, J., and Kuo, T., "A Comparative Analysis of Various EMG Pattern Recognition Methods", Medical Engineering Physics, 8, 5, 1996, pp 390-395.

Kohonen, T., "The Self-Organizing Map", Proceedings of the IEEE, 78, 1990, pp 1464-1480.

LaCourse, J.R., and Wilson, E.W., "BRAINIAC: A Brain-Computer Interface", Instrumentation and Measurement Society Newsletter, 1996, pp 9-14.

McMillan, G.R., Calhoun, G.L., Middendorf, M.S., Schnurer, J.H., Ingle, D.F., and Nasman, V.T., "Direct Brain Interface Utilizing Self-Regulation of the Steady-State Visual Evoked Response", in "Proceedings of the RESNA 18th Annual Conference", 1995, pp 693-695.

McMillan, G.R., Eggleston, R.G., and Anderson, T.R., "Nonconventional Controls", in Salvendy, G., (Ed) "Handbook of Human Factors and Ergonomics, 2nd Edition", New York, NY, Wiley, 1997, pp 729-771.

Nelson, W.T., Hettinger, L.J., Cunningham, J.A., Roe, M.M., Lu, L.G., Haas, M.W., Dennis, L.B., Pick, H.L., Junker, A., and Berg, C.B., "Brain-Body Actuated Control: Assessment of an Alternative Control Technology for Virtual Environments", in "Proceedings of the 1996 Image Conference", 1996, pp 225-232.

Paciga, J., Richard, P., and Scott, R.N., "Error Rate in Five-State Myoelectric Control Systems", Medical and Biological Engineering, 18, 1980, pp 287-290.

Parker, P., Scott, R.N., Hudgins, B., Hruczkowski, T., Hayden, J., and Englehart, K., "Coordinated / Simultaneous Control of a Multifunction Myoelectric Prosthesis", Progress Report #2 on NSERC Project No. CRD151174, April 1995.

Parker, P.A., and Scott, R.N., "Myoelectric Control of Prosthesis", CRC Critical Reviews in Biomedical Engineering, 13, Issue 4, 1986, pp 283-310.

Pfurtscheller, G., Flotzinger, D., and Neuper, C., "Differentiation between Finger, Toe and Tongue Movement in Man Based on 40 Hz EEG", Electroencephalography and Clinical Neurophysiology, 90, 1994, pp 456-460.

Pfurtscheller, G., Flotzinger, D., Mohl, W., and Peltoranta, M., "Prediction of the Side of Hand Movements from Single-Trial Multi-Channel EEG Data using Neural Networks", Electroencephalography and Clinical Neurophysiology, 82, 1992, pp 313-315.

Phillips, C.A., "Sensory Feedback Control of Upper- and Lower-Extremity Motor Prosthesis", CRC Critical Reviews in Biomedical Engineering, 16, 1988, pp 105-140.

Reiter, R., "Eine Neue Elektrounsthand", Grenzgebiete der Medizin, 4, 1948, pp 133-135.

Richard, P., Gander, R., Parker, P., and Scott, R.N., "Multistate Myoelectric Control: The Feasibility of 5-State Control", Journal of Rehabilitation Research and Development, 20, 1983, pp 84-86.

Saridis, G., and Stephanou, H., "A Hierarchical Approach to the Control of a Prosthetic Arm", IEEE Transactions on Systems Man and Cybernetics, SMC-7, 6, 1977, pp 407-420.

Saridis, G.N., and Gootee, T.P., "EMG Pattern Classification for a Prosthetic Arm", IEEE Transactions on Biomedical Engineering, BME-29, 6, 1982, pp 403-41.

Scott, R.N., Paciga, J., and Parker, P., "Operator Error in Multistate Myoelectric Control Systems", Medical and Biological Engineering, 16, 1978, pp 296-301.

Sullivan, G., Martell, C., Weltman, G., and Pierce, D., "Myoelectric Servo Control", Report to US Air Force Aeronautical Systems Division under Contract #AF33(657)-7771, May, 1963.

Sutter, E.E., "The Brain Response Interface: Communication through Visually-Induced Electrical Brain Responses", Journal of Microcomputer Applications, 15, 1992, pp 31-45.

Sutter, E.E., "The Visual Evoked Response as a Communication Channel", in "Proceedings: IEEE Symposium on Biosensors", IEEE, 1984, pp 95-100.

Taheri, B., Smith, R.L., and Knight, R.T., "A Dry Electrode for EEG recording", Electroencephalography and Clinical Neurophysiology, 90, 1994, pp 376-383.

United Kingdom Patent Number GB 2 274 396 B, December 1996.

Vatikiotis-Bateson, E., Munhall, K.G., Kasahara, Y., Garcia F., andYehia H., "Characterizing Audiovisual Information During Speech", in "Conference on Spoken Language Processing - CDROM", 1996, Paper No. 1010.

Vodovnik, L. "An Electromagnetic Brake Activated by Eyebrow Muscles", Electronics Engineering, 1967, pp 694-695.

Williams, T.W. "Practical Methods for Controlling Powered Upper-Extremity Prostheses", Assistive Technology, 2, 1, 1990, pp 3-18.

Wilson, G.F., and Eggemeier, T., "Psychophysiological Assessment of Workload in Multi-Task Environments", in Damos, D., (Ed) "Multiple Task Performance", Washington, DC, Taylor and Francis Press, 1991, pp 329-360.

Wilson, G.F., and Fisher, F., "Cognitive Task Classification Based upon Topographic EEG Data", Biological Psychology, 40, 1995, pp 239-250.

Wirta, R.W., Taylor, D.R., and Findley, F.R., "Pattern Recognition Arm Prosthesis: A Historical Perspective - Final Report", Bulletin of Prosthetics Research, 10-30, 1978, pp 9-35.

Wolpaw, J.R., and McFarland, D J., "Multichannel EEG-Based Brain-Computer Communication", Electroencephalography and Clinical Neurophysiology, 90, 1994, pp 444-449.

Wolpaw, J.R., McFarland, D.J., Neat, G.W., and Forneris, C.A., "An EEG-Based Brain-Computer Interface for Cursor Control", Electroencephalography and Clinical Neurophysiology, 78, 1991, pp 252-259.

7. INTEGRATION

"Safety Network to Detect Performance Degradation and Pilot Incapacitation" Papers presented at AMP Symposium. Tours, France. April 1990 AGARD-CP-490

Annett, J., Duncan, K.D., Stammers, R.B. and Gray, M.J., "Task Analysis", Training information Paper no. 6., London., HMSO, 1971.

Bailey, B., "Human Performance Engineering", Englewood Cliffs, NJ, Prentice Hall, 1982.

Barnard, P.J., and Hammond, N.V., "Cognitive contexts and interactive communication", IBM Hursley Human Factors Laboratory Report, 1983.

Bekey, G.A., "The human operator in control systems", in De Greene, K.B. (Ed) "Systems Psychology", New York, NY, McGraw Hill, 1970.

Boff, K.R., and Lincoln, J.E., "Engineering Data Compendium, Human Perception and Performance, Vols. I, II, and III", Armstrong Aerospace Medical Research laboratory, Ohio, 1988.

C.D. Wickens, "The structure of processing resources" in Nickerson, R. And Pew, R. (Eds) "Attention and Performance VIII", Hillsdale, NJ, Erlbaum.

CASHE:PVS, Version 1.0, Computer Aided Systems Human Engineering Performance Visualisation System. CSERIAC, Products Department, AL/CFH/CSERIAC Bldg 248, 2255 H Street, WPAFB OH 45433-7022, USA. (Email: cseriac@falcon.al.wpafb.af.mil).

Corker, K.M. and Smith, B.R., "An Architecture and model for cognitive engineering simulation analysis : Application to advanced aviation automation", in Proceedings of the 9th AIAA conference on Computing in Aerospace, New York. 1993, pp. 1079 - 1088.

Def Stan "Design and Airworthiness Requirements for Service Aircraft Vol Aeroplanes" AVP-970 published by MOD(PE) London (This is an evolving document in which Chapter 107 covers "Pilot's Cockpit - Controls and Instruments" and Chapter 105 covers "Crew Stations - General Requirements"

Dix, A., Finlay, J., Aboud, G. And Beale, R., "Human-Computer Interaction", Hemel Hempstead, UK: Prentice Hall, 1993.

Elkerton, J. And Williges, R.C., "Dialog design for intelligent interfaces", in Hancock, P.A., and Chignell, M.H. (Eds) "Intelligent Interfaces: Theory, Research and Design", Amsterdam, Elsevier, 1989.

Gawron, V.J., Anno, G., Fleishman, E.A., Jones, E.D., Lovesey, E.J., McGlynn, L.E., McMillan, G., McNally, R.E., Meister, D., O'Brien, L., Promisel, D.M., Ramirez, T., "Human Factors Taxonomy", in Proceedings of the Human Factors 35th Annual Meeting, 1991, pp. 1282-1287.

Hollnagel, E., and Cacciabue, P.C., "Cognitive Modelling in System Simulation", in Proceedings of the Third European Conference on Cognitive Science Approaches to Process Control, 2-6 September, Cardiff, UK, 1991.

Jarrett D. N Karavis A " Integrated flying helmets" Proc Instn Mech Engrs Vol 206, pp47-61 1992

Kieras, D.E., Towards a practical GOMS model methodology for user interface design", in Helander, M. (Ed) "Handbook of Human-Computer Interaction", Amsterdam, North Holland Elsevier, 1988, pp. 135-158.

Kieras, D.E., Wood, S.D., and Meyer, D.E., "Predictive Engineering Using the EPIC Architecture for a High Performance Task", in Proceedings of CHI'95, ACM, 1995.

Kirwan, B., "A Guide to Practical Human Reliability assessment", London, Taylor and Francis, 1994.

Kirwan, B., and Ainsworth, L.K., "A Guide to Task Analysis", London, Taylor & Francis, 1992.

McMillan, G.R., Eggleston, R.G., and Anderson, T.R., "Nonconventional Controls", in Salvendy, G. (Ed), "Handbook of Human factors and Ergonomics, 2nd Edition", New York, NY, Wiley, 1997.

Meister, D., "Behavioural analysis and measurement methods", New York, NY, Wiley, 1985.

MIL-STD-1776A (USAF) Aircrew Station and Passenger Accommodations, 1994

Moran, T.P., "The Command Language Grammar: A representation for the user interface of interactive computer systems", International Journal of Man Machine Studies, 15, 1981, pp. 3-50.

Rasmussen, J., Pejtersen, A.M. and Schmidt, K., "Taxonomy for Cognitive Work Analysis", in Proceedings of the First MOHAWK Workshop, Liege, May 15-16, 1990, Vol. 1 pp. 3-153.

Rencken, W.D., and Durrant-Whyte, H.F., "A quantitative model for adaptive task allocation in human-computer interfaces", in IEEE Transactions on Systems, Man and Cybernetics, 23 (4), 1993, pp. 1072-1090.

Ryder, J. And Zachary, W., "Experimental validation of attention switching component of the COGNET framework", in Proceedings of the Human Factors Society 35th Annual Meeting, 1991.

Schneiderman, B., "Designing the User Interface: Strategies for Effective Human-Computer Interaction", Massachusetts, Addison Wesley, 1987.

Shepherd, A., "Hierarchical task Analysis and Training Decisions", Programmed Learning and Educational Technology, 22, pp. 162-176, 1985.

Shoval, S., Koren, Y., and Borenstein, J., "Optimal task allocation in task agent control space", in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 4, 1993, pp. 27-32.

Smith, S.L. and Mosier, J.N., "Design guidelines for user-system interface software", The Mitre Corporation Technical Report ESD-TR-84-358, 1984.

Swain, A.D., and Guttman, G., "Handbook for Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications", Report NUREG/CR-1278, US Nuclear Regulatory Commission, Washington, DC, 1983.

van Someren, M.W., Barnard, Y.F., and Sandberg, J.A.C., "The Think Aloud Method, A Practical Guide to Modelling Cognitive Processes", London, Academic Press, 1994.

Warren, C.P., Day, P.O., Hook, M.K. and Hicks, M., "Performance and Usability Modelling in Air traffic control (PUMA)", in Proceedings of the Fourth International Conference on Human-Machine Interaction and Artificial Intelligence in Aerospace, Toulouse, Sept. 28-30, 1993.

Williges, R.C., Williges, B. And Elkerton, J., "Software Interface Design", in Salvendy, G., (Ed) "Handbook of Human Factors", New York, NY, Wiley, 1987.

8. POTENTIAL APPLICATIONS

Chapman, D.D., and Simmons, J.R., "A Comparative Evaluation of Voice Versus Keypad Input For Manipulating Electronic Technical Data For Flight Line Maintenance Technicians", M.S. Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, USA, Sept. 1995.

Lèger A., Portier, L., Badou, J., and Trosseille, X. "Crash Survivability and Operational Comfort Issues of Helmet Mounted Displays in Helicopters: Simulation Approach and Flight Test Results." Communication Seminar on Helmet Mounted Design, Framingham, Ma. December 1997.

McEntire, B.J., Shanahan, D.F. "Mass Requirements for Helicopter Aircrew Helmets." AGARD CP597. 1997

Mertz, H.J, "Anthropomorphic Test Devices" in Nahum, A.M. & Melvin, J.M. (eds) "Accidental Injury: Biomechanics and Prevention" Springer-Verlag, New York, 1993

Tatham, N.O. The effect of turbulence on Helmet Mounted sight aiming accuracies AGARD CPP-267 High Speed Low Level Flight: Aircrew Factors 1979

9. MULTIMODALITY

Abowd G., Bowen J., Dix A., Harrison M., Took R., « User interfaces languages : a survey of existing methods », Rapport Tech. PRG-TR-5-89, Oxford University, 1989.

Bolt R. A., « Put that there : voice and gesture at the graphic interface », Computer Graphics, 14 pp. 262-270, 1980.

Bresnan J., Kaplan R., « Lexical functional grammars ; a formal system for grammatical representation », in 'The mental representation of grammatical relations', Bresnan J. (éd.), MIT Press, Cambridge, pp. 173-281, 1982.

Brun P., « XTL : a temporal logic for the formal development of interactive systems », in 'Formal methods in human-computer interaction', Philippe Palanque et Fabio Paterno (éds.), Springer- Verlag (Londres), pp. 121-140, 1998.

Brun P., Beaudouin-Lafon M., « A taxonomy and evaluation of formalisms for the specification of interactive systems », in Proc. of 10th annual conference of the British Human-Computer Interaction Group (HCI'95), University of Huddersfeild, UK, août-septembre 1995.

Cadoz C., « Le retour d'effort dans la communication gestuelle avec la machine - Le concept de communication instrumentale », Actes des 3 èmes journées Internationales de Montpellier, L'interface des mondes réels et virtuels, 7-11 février 1994.

Coutaz J., Duke D., Faconti G., Harrison M., Paterno F., « Formal methods and multimodal interactive systems », Rapport Tech. sm/wp61, ESPRIT BRA 7040 Amodeus-2, septembre 1995, http://www.mrc-apu.cam.ac.uk/amodeus/

Coutaz J., Nigay L., Salber D., Blandford A., May J., Young R. M., « Four easy pieces for assesing the usability of multimodal interaction : the CARE properties », Proc. INTERACT'95, Lillehammer, Norvège, juin 1995.

Dauchy P., Mignot C., Valot C., « Joint speech and gesture analysis. Some experimental results on multimodal interface », EUROSPEECH'93, Berlin.

Duce D.A., Duke D.J., « Syndetic modelling : a new opportunity for formal methods », Rapports Tech. ID/WP57, ESPRIT BRA 7040 Amodeus-2, décembre 1992, http://www.mrc-apu.cam.ac.uk/amodeus/

Duke D.J., Harrison M.D., «'Synergistic' interactors », Rapport Tech. SM/WP03, ESPRIT BRA 7040 Amodeus-2, décembre 1992, http://www.mrc-apu.cam.ac.uk/amodeus/

Duke D.J., Harrison M.D., « Abstract interaction objects », Comp. Graphics Forum, 12(3), pp.25-36, 1993.

Gaiffe B., Romary L., Pierrel J.M., « References of a multimodal dialogue : towards a unified processing », EUROSPEECH'91, 2nd European Conf. on Speech Communication and Technology, vol. 3, pp. 1481-1485, 1991.

Harrison M.D., Abowd G.D., Dix A.J., « Analysing display oriented interaction by means of system models », in 'Computers, communication and usability : design issues, research and methods for integrated services', pp. 147-163, Elsevier, 1993.

Koenderink J.J., « The concept of local sign », in 'Limits in Perception', A.J. van Doorn et al. (éds.), VNU Science Press, pp. 495-547, 1984.

Martin J.C., « Cadre d'étude de la multimodalité fondé sur des types et buts de coopération entre modalités », Actes des 3èmes Journées Internationales de Montpellier, L'interface des mondes réels et virtuels, 7-11 février 1994.

Nielsen J., « A virtual protocol model for computer-human interaction », Int. J. of Man-Machine Studies, 24, pp. 301-312, 1986.

Nigay L., Coutaz J., « A design space for multimodal interfaces : concurrent processing and data fusion «, in Proc. INTERCHI'93 Human Factors in Computing Systems, Amsterdam, 24-29 avril 1993, ACM Press, pp. 172-178.

Nigay L., Coutaz J., « A generic platform for adressing the multimodal challenge », Proc. CHI'95, Denver, mai 1995.

Nigay L., Coutaz J., « Software architecture modelling : bridging two worlds using ergonomics and software properties », in 'Formal methods in human-computer interaction', Philippe Palanque et Fabio Paterno (éds.), Springer-Verlag (Londres), pp. 49-74, 1998.

Roast C.R., Harrison M.D., « The specification and prototyping of interaction models using dynamic logic », HCI Group, University of York, avril 1989.

Taylor M.M., « Layered protocols for computer-human dialogue. I : Principles », Int. J. Man-Machine Studies (1988), 28, pp. 175-218.

Teil D., Ferrari S., « Communication multimodale - Approche ascendante, application au domaine spatial », Ecole SIC (Systèmes d'Information et de Communication) '94, Campus thomson, 24 avril 1994.

Thimbleby H., « Generative user-engineering principles for user interface design » in Proc. HCI - INTERACT'84, B. Shackel (éd.), Elsevier Science Publishers (North-Holand), 1985.

Tylor M.M., « Layered protocols for computer-human dialogue. II : Some practical issues », Int. J. Man-Machine Studies (1988), 28, pp. 219-257.

10. WAVELETS

Daubechies I., "Ten Lectures on Wavelets", Society for Industrial and Applied Mathematics, 1992, Philadelphia

Kronland-Martinet.R. "The Wavelet Transform for Analysis, Synthesis and Processing of Speech and Music Sounds" Computer Music Journal, Vol 12 No. 4 1988.

Mallat S., "A theory for Multi-Resolution Signal Decomposition: the Wavelet Representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, no. 7, 1989.

Wickerhauser V., "Adapted Wavelet Analysis from Theory to Software", A. K. Peters, 1994, Massachusetts.

11. DYNAMIC TIME WARPING

Viterberi, A. J. « Error bounds for convolution codes and an asymptotically optimal decoding algorithm. » IEEE Trans. on Information Theory, IT-13 1967 pp 260-269

12. NEURAL NETWORKS

Hinton, G.E. "Connectionist learning procedures", Artificial Intelligence, Vol. 40, 1989, pp 185-234.

Hush, D.R. and Horne, B.G. "Progress in Supervised Neural Networks, IEEE Signal Processing Mag., January 1993, pp 8-39.

Lippman, R.P. "An introduction to computing with neural nets", IEEE ASSP Mag. April, 1987, pp 4-22.

Lippman, R.P. "Pattern classification using neural networks", IEEE Communications Mag., November, 1989, pp. 47-64.

Pao, Y.H. "Adaptive Pattern Recognition and Neural Networks", Addison-Wesley, Reading, Mass., 1988.

Rumelhart, D.E. and McClelland, J.L. "Parallel Distributed Processing:Explorations in the Microstructure of Cognition", (Vol. I), Cambridge, Ma., MIT Press, 1986, (ISBN 0-262-18120-7).

Viterbi, A. J. "Error bounds for convolution codes and an asymptotically optimal decoding algorithm." IEEE Trans. on Information Theory, IT-13 1967 pp 260-269

Werbos, P. "Beyond regression: New tools for prediction and analysis in the behavioural sciences", PhD dissertation, Harvard University, 1974.

13. HIDDEN MARKOV MODELS

Bahl, L., Jelinek, F. and Mercer, R., "A maximum likelihood approach to continuous speech recognition", IEEE Trans. On Pattern Analysis and Machine Inteligence, PAMI-5(2), 1983, pp 179-190.

Baker, J., "Stochastic Modeling for automatic speech understanding", in R. Reddy (ed), "Speech Recognition", Academic Press, 1975, pp 521-542.

Baum, L., "An inequality and associated maximization technique in statistical estimation for probabilistic functions for Markov processes", Inequalities, No. 3, 1972, pp 1-8.

Forney, G. D., "The Viterbi Algorithm", Proc. IEEE, Vol. 61, 1973, pp 268-278.

Jelinek, F., "Continuous speech recognition using statistical methods", Proc. IEEE, 64(4), 1976, pp 532-556.

Lee, K.-F., "Large-Vocabulary Speaker-Independent Continuous Speech Recognition: The SPHINX System", Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA, 1988.

Levinson, S., Rabiner, L., and Sondhi, M., "An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition", Bell Sys. Tech. J., 62(4), 1983, pp 1035-1074.

Rabiner, L., "A tutorial on hidden Markov models and selected applications in speech recognition", Proc. IEEE, 77(2), 1989, pp 257-286.

Rabiner, L., and Juang, B.-W., "An introduction to hidden Markov models", IEEE ASSP Magazine, 3(1), 1986, pp 4-16.

Rabiner, L., and Juang, B.-W., "Fundamentals of Speech Recognition", Englewood Cliffs, NJ, Prentice-Hall, 1993. (ISBN 0 13 015157 2)

Rabiner, L., and Juang, B.-W., "Hidden Markov models for speech recognition - strengths and limitations", in P. Laface and R. De Mori, (Eds) "Speech Recognition and Understanding. Recent Advances, Trends and Applications", Springer-Verlag, 1992, pp 3-29.

Schwartz, R., and Kubala, F., "Hidden Markov models and speaker adaptation", in P. Laface and R. De Mori, (Eds) "Speech Recognition and Understanding. Recent Advances, Trends and Applications", Springer-Verlag, 1992, pp 31-57.

Viterbi, A., "Error bounds for convolutional codes and an asymptotically optimal decoding algorithm", IEEE Trans. On Information Theory, IT-13, 1967, pp 260-269.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	RTO-EN-3 AC/323(HFM)TP/1	ISBN 92-837-1003-7	UNCLASSIFIED/ UNLIMITED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
BP 25, 7 rue Ancelle, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Alternative Control Technologies: Human Factors Issues

7. Presented at/sponsored by

The material in this publication was assembled to support a Lecture Series under the sponsorship of the Human Factors and Medecine Panel and the Consultant and Exchange Programme of RTO presented on 7-8 October 1998 in Brétigny, France, and on 14-15 October 1998 at Wright-Patterson Air Force Base, Ohio, USA.

8. Author(s)/Editor(s)	9. Date
Multiple	October 1998

10. Author's/Editor's Address	11. Pages
Multiple	116

12. Distribution Statement There are no restrictions on the distribution of this document. Information about the availability of this and other RTO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Human factors engineering
Artificial intelligence
Man machine systems
Control equipment
Speech recognition
Cockpits
Design
Adaptive systems
Automatic control
Integrated systems
Pilots (personnel)
Motion studies
Eye movements
Man computer interface
Computerized simulation
Voice communication
Manual controls
Head (anatomy)
Eye (anatomy)
Tracking (position)

14. Abstract

With the increasing intelligence of computer systems, it is becoming more desirable to have an operator communicate with machines rather than simply operate them. In combat aircraft, this need to communicate is made quite crucial due to high temporal pressure and workload during critical phases of the flight (ingress, engagement, deployment of self-defence). The HOTAS concept, with manual controls fitted on the stick and throttle, has been widely used in modern fighters such as F16, F18, EFA and Rafale. This concept allows pilots to input real time commands to the aircraft system. However, it increases the complexity of the pilot task due to inflation of real time controls, with some controls being multifunction. It is therefore desirable, in the framework of "ecological interfaces", to introduce alternative input channels in order to reduce the complexity of manual control in the HOTAS concept and allow more direct and natural access to the aircraft systems.

Control and display technologies are the critical enablers for these advanced interfaces. There are a variety of novel alternative control technologies that when integrated usefully with critical mission tasks can make natural use of the innate potential of human sensory and motor systems. Careful design and integration of candidate control technologies will result in human-machine interfaces which are natural, easier to learn, easier to use, and less prone to error. Significant progress is being made on using signals from the brain, muscles, voice, lip, head position, eye position and gestures for the control of computers and other devices.

Judicious application of alternative control technologies has the potential to increase the bandwidth of operator-system interaction, improve the effectiveness of military systems, and realise cost savings. Alternative controls can reduce workload and improve efficiency within the cockpit, directly supporting the warfighter.

By the end of 1997, WG 25 had extensively reviewed human factor aspects of current and prospective alternative technologies along with operational needs and integration issues. Dissemination of the knowledge among Engineering and Human Factor communities has to be made as early as possible to facilitate implementation of these new technologies in future projects.

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25 • 7 RUE ANCELLE

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Télécopie 0(1)55.61.22.99 • Télex 610 176

DIFFUSION DES PUBLICATIONS

RTO NON CLASSIFIEES

L'Organisation pour la recherche et la technologie de l'OTAN (RTO), détient un stock limité de certaines de ses publications récentes, ainsi que de celles de l'ancien AGARD (Groupe consultatif pour la recherche et les réalisations aérospatiales de l'OTAN). Celles-ci pourront éventuellement être obtenues sous forme de copie papier. Pour de plus amples renseignements concernant l'achat de ces ouvrages, adressez-vous par lettre ou par télécopie à l'adresse indiquée ci-dessus. Veuillez ne pas téléphoner.

Des exemplaires supplémentaires peuvent parfois être obtenus auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus sur la liste d'envoi de l'un de ces centres.

Les publications de la RTO et de l'AGARD sont en vente auprès des agences de vente indiquées ci-dessous, sous forme de photocopie ou de microfiche. Certains originaux peuvent également être obtenus auprès de CASI.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, B-1140 Bruxelles

CANADA
Directeur - Gestion de l'information
(Recherche et développement) - DRDGI 3
Ministère de la Défense nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

ETATS-UNIS
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076

FRANCE
O.N.E.R.A. (Direction)
29, Avenue de la Division Leclerc
92322 Châtillon Cedex

GRECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Aeronautica Militare
Ufficio Stralcio RTO/AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PAYS-BAS
RTO Coordination Office
National Aerospace Laboratory NLR
P.O. Box 90502
1006 BM Amsterdam

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

ROYAUME-UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

AGENCES DE VENTE

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076
Etats-Unis

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume-Uni

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Document Delivery,
Montreal Road, Building M-55
Ottawa K1A 0S2
Canada

Les demandes de documents RTO ou AGARD doivent comporter la dénomination "RTO" ou "AGARD" selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants:

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de ressources uniformes (URL) suivant:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
Etats-Unis

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
Etats-Unis
(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

Imprimé par le Groupe Communication Canada Inc.
(membre de la Corporation St-Joseph)
45, boul. Sacré-Cœur, Hull (Québec), Canada K1A 0S7

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION
BP 25 • 7 RUE ANCELLE
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Telefax 0(1)55.61.22.99 • Telex 610 176

DISTRIBUTION OF UNCLASSIFIED RTO PUBLICATIONS

NATO's Research and Technology Organization (RTO) holds limited quantities of some of its recent publications and those of the former AGARD (Advisory Group for Aerospace Research & Development of NATO), and these may be available for purchase in hard copy form. For more information, write or send a telefax to the address given above. **Please do not telephone.**

Further copies are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO publications, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your organisation) in their distribution.

RTO and AGARD publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form. Original copies of some publications may be available from CASI.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, B-1140 Bruxelles

CANADA
Director Research & Development
Information Management - DRDIM 3
Dept of National Defence
Ottawa, Ontario K1A 0K2

DENMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92322 Châtillon Cedex

GERMANY
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio Stralcio RTO/AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
See Belgium

NETHERLANDS
RTO Coordination Office
National Aerospace Laboratory, NLR
P.O. Box 90502
1006 BM Amsterdam

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

SPAIN
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

TURKEY
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076
United States

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Document Delivery,
Montreal Road, Building M-55
Ottawa K1A 0S2
Canada

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform resource locator:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
United States

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 22161
United States
(also available online in the NTIS Bibliographic Database or on CD-ROM)

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

ISBN 92-837-1003-7